

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

THIS PAGE BLANK (USPTO)

77515

Delaval, Jan

From: Roark, Jessica
Sent: Wednesday, October 09, 2002 8:49 AM
To: Delaval, Jan
Subject: 09/806955

Jan,

Please search, including pending, the following from 09/806955:

SEQ ID NO:1 and
SEQ ID NO:2.

Results on paper please.

Thanks!

Jessica H. Roark

CM1 8A03
Mailbox 9E12
Art Unit 1644
703 605-1209

revised
*SEQ 2 = SEQ 1 except
SEQ 1 has a C-terminus 6 His tag!*

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:10:24 ; Search time 26.02 seconds

(without alignments)
599.845 Million cell updates/sec

Title: US-09-806-955a-1

Perfect score: 3273

Sequence: 1 MEDKKEDVGVVGLDCTT.....AGPPPGEDTALHHHHH 639

Scoring table:

BLOSUM62
Gapco 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

1 number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3209	98.0	654	1	US-08-441-139-11 Sequence 11, Appl
2	3168	96.8	666	1	US-08-441-139-16 Sequence 16, Appl
3	2174	66.4	682	1	US-08-441-139-2 Sequence 2, Appl
4	2156	65.9	663	1	US-08-441-139-7 Sequence 7, Appl
5	2077	63.5	646	1	US-08-441-139-14 Sequence 14, Appl
6	2065.5	63.1	679	1	US-08-441-139-5 Sequence 5, Appl
7	1938	59.2	643	4	US-08-797-358B-3 Sequence 3, Appl
8	1547.5	47.3	679	1	US-08-214-583-2 Sequence 2, Appl
9	1474	45.0	641	1	US-08-441-139-4 Sequence 4, Appl
10	1461.5	44.7	649	4	US-09-066-047-5 Sequence 5, Appl
11	1444	44.1	607	2	US-08-472-534-5 Sequence 5, Appl
12	1303.5	39.8	600	6	5240706-1 Patent No. 5240706
13	1140.5	34.8	339	2	US-08-928-692-52 Sequence 52, Appl
14	941	28.8	187	6	5196523-13 Patent No. 5196523
15	824	25.2	168	1	US-08-441-139-10 Sequence 10, Appl
16	818.5	25.0	315	1	US-08-257-073-7 Sequence 7, Appl
17	750.5	22.9	471	1	US-08-203-905B-2 Sequence 2, Appl
18	726.5	22.2	472	1	US-08-203-905B-14 Sequence 14, Appl
19	701	21.4	307	4	US-08-858-207A-481 Sequence 481, App
20	643.5	19.6	129	6	5196523-10 Patent No. 5196523
21	607.5	18.6	999	2	US-08-770-301A-3 Sequence 3, Appl
22	607.5	18.6	999	3	US-09-175-581-3 Sequence 3, Appl
23	598	18.3	999	2	US-08-770-301A-1 Sequence 1, Appl
24	598	18.3	999	3	US-09-175-581-1 Sequence 1, Appl
25	528.5	16.1	560	2	US-08-928-692-53 Sequence 53, Appl
26	437.5	13.4	374	2	US-08-928-692-51 Patent No. 5196523
27	402	12.3	77	6	5196523-7

28	395	12.1	79	6	5196523-11	Patent No. 5196523
29	367.5	11.2	599	4	US-09-080-983-9	Sequence 9, Appl
30	358	10.9	549	2	US-08-770-544-6	Sequence 6, Appl
31	345.5	10.6	80	1	US-08-464-164-4	Sequence 4, Appl
32	345.5	10.6	80	1	US-08-338-057-4	Sequence 4, Appl
33	345.5	10.6	80	2	US-08-668-416-4	Sequence 4, Appl
34	323	9.9	554	5	PCT-US94-06430-7	Sequence 7, Appl
35	320	9.8	554	4	US-08-591-468-7	Sequence 7, Appl
36	267.5	8.2	714	2	US-08-472-934-3	Sequence 3, Appl
37	264	8.1	56	6	5196523-12	Patent No. 5196523
38	230	7.0	46	6	5196523-8	Patent No. 5196523
39	211	6.4	42	6	5196523-15	Patent No. 5196523
40	192	5.9	38	6	5196523-9	Patent No. 5196523
41	168.5	5.1	1786	4	US-08-973-462-8	Sequence 8, Appl
42	166	5.1	125	4	US-08-858-207A-407	Sequence 407, App
43	151.5	4.6	1805	1	US-07-853-913-2	Sequence 2, Appl
44	147	4.5	693	4	US-08-235-836C-68	Sequence 68, Appl
45	142.5	4.4	2283	4	US-09-308-373-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-441-139-11
Sequence 11, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: Wiltrop, Dr. Karl D.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: DIGILIO, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-11
Query Match 98.0%, Score 3209; DB 1; Length 654;
Best local similarity 99.7%, Pred. No. 5.2e-250;
Matches 629; Conservative 0; Indels 0; Gaps 0;

QY 2 DEEKKEDVGTGTVGIDLGTTTSCVGFKNRGVEIITANDGNRITPSYVAFPEGEERLIGDA 61
DB 20 DEEKKEDVGTGTVGIDLGTTTSCVGFKNRGVEIITANDGNRITPSYVAFPEGEERLIGDA 79
QY 62 AKNOLTSNPENTVFDKRLIGRTWDPVSVOODIKFLPKVYEKTKPYIOVDIGGQTKT 121
DB 80 AKNOLTSNPENTVFDKRLIGRTWDPVSVOODIKFLPKVYEKTKPYIOVDIGGQTKT 139
QY 122 FAPBEISAMVLTAKMKTABAYLGKVTNAVVPVAFPDORQATKDACTIAGLNVKRII 181
DB 140 FAPBEISAMVLTAKMKTABAYLGKVTNAVVPVAFPDORQATKDACTIAGLNVKRII 199
QY 182 NEPTAAIAVGLDKRGEKNILVFDLGGCTFDVSLITIDNCVFEVATNGDTHLGGEDPD 241
DB 200 NEPTAAIAVGLDKRGEKNILVFDLGGCTFDVSLITIDNCVFEVATNGDTHLGGEDPD 259
QY 242 QRVMEHFIRLKYKKTGKDVKNRAVOKLRREVEKAKRALSQHOARIEIESFEGBEDFS 301
DB 260 QRVMEHFIRLKYKKTGKDVKNRAVOKLRREVEKAKRALSQHOARIEIESFEGBEDFS 319
QY 302 ETLTRAKFEELNMDLFRSTMKPVOKVLEDSDLKSDIDEIVLVGSGSTRIPKIQOLVKEFF 361
DB 320 ETLTRAKFEELNMDLFRSTMKPVOKVLEDSDLKSDIDEIVLVGSGSTRIPKIQOLVKEFF 379
QY 362 NGKEPSRGINPDEAVAYGAAGVLSGDDPTGDLVLLDVCPRTLGIETVGVMTKLIIPR 421
DB 380 NGKEPSRGINPDEAVAYGAAGVLSGDDPTGDLVLLDVCPRTLGIETVGVMTKLIIPR 439
QY 422 NTVPYTKKSQIFSTASDNOPVTYIKVYEGSERPLTKDNHLGTFDLTGIPAPRGVPOIEV 481
DB 440 NTVPYTKKSQIFSTASDNOPVTYIKVYEGSERPLTKDNHLGTFDLTGIPAPRGVPOIEV 499
QY 482 TFEIDVNGILRTVAEDKGTGNKNKITINDONRLPPEIERMVNDARKFAEEDKKLKERI 541
DB 500 TFEIDVNGILRTVAEDKGTGNKNKITINDONRLPPEIERMVNDARKFAEEDKKLKERI 559
QY 542 DTRNELESYAYSILKNOIGDKERLGSSEDEKEMEKAVEKTIEMLSHODADIEDPKAK 601
DB 560 DTRNELESYAYSILKNOIGDKERLGSSEDEKEMEKAVEKTIEMLSHODADIEDPKAK 619
QY 602 KKELEEVOPITISKLYGSAGPPPTGEEDTAE 632
DB 620 KKELEEVOPITISKLYGSAGPPPTGEEDTAE 650

RESULT 2
US-08-441-139-16
Sequence 16, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: Wiltcup, Dr. Karl D.
APPLICANT: Robinson, Anne S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-16

Query Match 96.8%; Score 3168; DB 1; Length 666;
Best Local Similarity 97.8%; Pred. No. 1,1e-246;
Matches 617; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

QY 2 DEEKKEDVGTGTVGIDLGTTTSCVGFKNRGVEIITANDGNRITPSYVAFPEGEERLIGDA 61
DB 32 DEEKKEDVGTGTVGIDLGTTTSCVGFKNRGVEIITANDGNRITPSYVAFPEGEERLIGDA 91
QY 62 AKNOLTSNPENTVFDKRLIGRTWDPVSVOODIKFLPKVYEKTKPYIOVDIGGQTKT 121
DB 92 AKNOLTSNPENTVFDKRLIGRTWDPVSVOODIKFLPKVYEKTKPYIOVDIGGQTKT 151
QY 122 FAPBEISAMVLTAKMKTABAYLGKVTNAVVPVAFPDORQATKDACTIAGLNVKRII 181
DB 152 FAPBEISAMVLTAKMKTABAYLGKVTNAVVPVAFPDORQATKDACTIAGLNVKRII 211
QY 182 NEPTAAIAVGLDKRGEKNILVFDLGGCTFDVSLITIDNCVFEVATNGDTHLGGEDPD 241
DB 212 NEPTAAIAVGLDKRGEKNILVFDLGGCTFDVSLITIDNCVFEVATNGDTHLGGEDPD 271
QY 242 QRVMEHFIRLKYKKTGKDVKNRAVOKLRREVEKAKRALSQHOARIEIESFEGBEDFS 301
DB 272 QRVMEHFIRLKYKKTGKDVKNRAVOKLRREVEKAKRALSQHOARIEIESFEGBEDFS 331
QY 302 ETLTRAKFEELNMDLFRSTMKPVOKVLEDSDLKSDIDEIVLVGSGSTRIPKIQOLVKEFF 361
DB 332 ETLTRAKFEELNMDLFRSTMKPVOKVLEDSDLKSDIDEIVLVGSGSTRIPKIQOLVKEFF 391
QY 362 NGKEPSRGINPDEAVAYGAAGVLSGDDPTGDLVLLDVCPRTLGIETVGVMTKLIIPR 421
DB 392 NGKEPSRGINPDEAVAYGAAGVLSGDDPTGDLVLLDVCPRTLGIETVGVMTKLIIPR 451
QY 422 NTVPYTKKSQIFSTASDNOPVTYIKVYEGSERPLTKDNHLGTFDLTGIPAPRGVPOIEV 481
DB 452 NTVPYTKKSQIFSTASDNOPVTYIKVYEGSERPLTKDNHLGTFDLTGIPAPRGVPOIEV 511
QY 482 TFEIDVNGILRTVAEDKGTGNKNKITINDONRLPPEIERMVNDARKFAEEDKKLKERI 541
DB 512 TFEIDVNGILRTVAEDKGTGNKNKITINDONRLPPEIERMVNDARKFAEEDKKLKERI 571
QY 542 DTRNELESYAYSILKNOIGDKERLGSSEDEKEMEKAVEKTIEMLSHODADIEDPKAK 601
DB 572 DTRNELESYAYSILKNOIGDKERLGSSEDEKEMEKAVEKTIEMLSHODADIEDPKAK 631
QY 602 KKELEEVOPITISKLYGSAGPPPTGEEDTAE 632
DB 632 KKELEEVOPITISKLYGSAGPPPTGEEDTAE 662

RESULT 3
US-08-441-139-2
Sequence 2, Application US/08441139
Patent No. 5773245


```

||||| ||||||| ||||||| :||| ||| ||| :||| :|||
Db AKNOAPSPPENTIFDIKRLGKFKDEKTMAKIKSPFHIVNDKRRPLVEKNNV-GCKKKK 144
Qy 122 FAPETISAMVLTMMKEPAEAYLGCKKTAHVTVPAVFNDQROATKAGTAGLVNKRIT 181
Db 145 FTPEISAMVLTMMKEPAEAYLGCKKTAHVTVPAVFNDQROATKAGTAGLVNKRIT 204
Qy 182 NEPTAALAVGLCKREGEKILVFDLGGCTFPVSLTIDNGVEFVAVANGDTHLGGDFD 241
Db 205 NEPTAALAVGLCKREGEKILVFDLGGCTFPVSLTIDNGVEFVAVANGDTHLGGDFD 264
Qy 242 QRVMEHFKIKYKKKTKGDKVRKNRANOKLRREVAKARALSSQHOARLEISFEGEDFS 301
Db 265 NRVNIVLARYKRNKNVNDYTKDKLAKGKLRREVAKANGTSSQKSVRLISFNGDGS 324
Qy 302 ETLTRAFEEELNMDLFRSTMKPVQKVLIEDSDLKSDIDEIVLVGSTRIRPKIQOLVKEFP 361
Db 325 ETLTRAFEEELKHSIQDEPEVEYQVLKDSNLKKESEIDIVLVGSTRIRPKVQGLLESFF 384
Qy 362 NGKPSRGINPDEAVAGAAVQGVLSGDQDGLVLDVCPRLTIGITGVGVTKLIPR 421
Db 385 -GKASGIPNDEAVAGAAVQGVLSGEGSDNVLVDVLPRLTIGITGVGVTKLIGR 443
Qy 422 NTVPTRKSSQIFSTASDNOPTVIKVEGGERPLTKDNHLGTFDLTGIPAPRGVPOIEV 481
Db 444 NTPTRKSSQIFSTASDNOPTVIKVEGGERPLTKDNHLGTFDLTGIPAPRGVPOIEV 503
Qy 482 TFEIDVNGILRYTAEDK-GTGNKNTKITINDONRLTPEIEIRNVADAKKFEEDKKLIPR 540
Db 504 TFEIDVNGILRYTAEDK-GTGNKNTKITINDONRLTPEIEIRNVADAKKFEEDKKLIPR 563
Qy 541 IDTRNELESYAVSLKNOIGDKERLGSSEDEKETMEKAVEKEITMLESN-QDADIEDFK 599
Db 564 IEANNTLENTYAVSLKNOIGDKERLGSSEDEKETMEKAVEKEITMLESN-QDADIEDFK 623
Qy 600 AKKELEIYQPIISKLVSAGSPPPTG 629
Db 624 DQROKLDVAVHPITQKLY-SEGAGDADEED 652

```

```

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELER: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 646 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-441-139-14

Query Match 63.5%; Score 2077; DB 1; Length 646;
Best Local Similarity 66.0%; Pred. No. 6-5e-159;
Matches 409; Conservative 97; Mismatches 108; Indels 6; Gaps 5;

Qy 10 GTAVGIDGTTSCGVGFKNGREYIITANOGRNTPSYAFPEDEBRILGAAKNOGLSN 69
Db 4 GRAVGIDGTTSCGVGFKNGREYIITANOGRNTPSYAFPEDEBRILGAAKNOGLSN 62
Qy 70 PENTVFDAKRLIGRTWMDPSVQODIFLPFKVYEKKTKPYIYVDIGGQTKTFAPAEISA 129
Db 63 PNTVFDAKRLIGRFDVAVYQSDMKHMPVNVNDGRPKVOYEV-KGETKSFYDEEVS 121
Qy 130 MVLTRKMETAEAYLGCKKTAHVTVPAVFNDQROATKAGTAGLVNKRITNEPTAAI 189
Db 122 MVLTRKMETAEAYLGCKKTAHVTVPAVFNDQROATKAGTAGLVNKRITNEPTAAI 181
Qy 190 AYGLDKREG-EKNILVFDLGGCTFPVSLTIDNGVEFVAVANGDTHLGGDFDQRVMEH 248
Db 182 AYGLDKREG-EKNILVFDLGGCTFPVSLTIDNGVEFVAVANGDTHLGGDFDQRVMEH 241
Qy 249 IKLYKKTKGKVRDNRAVOKLRREVAKARALSSQHOARLEISFEGEDSETLTPRAK 308
Db 242 IAEKRRKKKNDISNKRRAVRLTACERKRLTSSSTQSIIDSLYEGIDFYISITPAR 301
Qy 309 FEELNMDLFRSTMKPVQKVLIEDSDLKSDIDEIVLVGSTRIRPKIQOLVKEFFNGKPEPS 368
Db 302 FEELNMDLFRSTMKPVQKVLIEDSDLKSDIDEIVLVGSTRIRPKIQOLVKEFFNGKPEPS 361
Qy 369 GINDEAVAYGAQAQVAVLSGD--QPTGDLVLDVCPRLTIGITGVGVTKLIPR 426
Db 362 SINDEAVAYGAQAQVAVLSGDSENVQDILLDVPRLSLGIEFGVGVTKLIPR 421
Qy 427 TKKSQISTASDNOPTVIKVEGGERPLTKDNHLGTFDLTGIPAPRGVPOIEVTEPID 486
Db 422 TKQVQITFTYSNDQPGVLIQVYGEGRAMKRNKLLGKFEELTGIPAPRGVPOIEVTEPID 481
Qy 487 VNGILRYTAEDKGTGNKNTKITINDONRLTPEIEIRNVADAKKFEEDKKLKERIDTRNE 546
Db 482 ANGLINVASAVKSGKKNKLTITNDKGRLSKDEIMVQEAKEYAEDKQRODKVSSNS 541
Qy 547 LESYAVSLKNOIGDKERLGSSEDEKETMEKAVEKEITMLESNODADIEDFKAKKLE 606
Db 542 LESYAVSLKNOIGDKERLGSSEDEKETMEKAVEKEITMLESNODADIEDFKAKKLE 600
Qy 607 EIVQPIISKLVSAGSPPPTG 626
Db 601 KVCNPIITKLVSAGSGMPCG 620

RESULT 6
US-08-441-139-5
; Sequence 5, Application US/08441139
; Patent No. 5773245
; GENERAL INFORMATION:
; APPLICANT: Wiltup, Dr. Karl D.
; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESSES:

```


ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
 STREET: 400 Garden City Plaza
 CITY: Garden City
 STATE: NY
 COUNTRY: USA
 ZIP: 11530
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/441,139
 FILING DATE: 15-MAY-1995
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/089,997
 FILING DATE: 06-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Digilio, Frank S.
 REGISTRATION NUMBER: 31,346
 REFERENCE/DOCKET NUMBER: 8646
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 516-742-4343
 TELEFAX: 516-742-4366
 TELEX: 230 901 SANS UR
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 679 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-441-139-5

Query Match 63.1%; Score 2065.5; DB 1; Length 679;
 Best Local Similarity 65.4%; Pred. No. 5.9e-158;
 Matches 403; Conservative 90; Mismatches 118; Indels 5; Gaps 4;

QY 7 EDVGTAVGIDLTGTTSCVGVFNKGRVEILANDQGRITPSVATPGEERLIGDAKNOL 66
 DB 48 EDGTVIGIDLTGTTSCVAVMKNKTEILANDQGRITPSVSEFT-DDEERLIGDAKNQA 106
 QY 67 TSNPEVTVEDAKRLIGRTWNDSPVOODIKFLPFKYVEKRTPTIYQVIGGQGTTFAPAE 126
 DB 107 ASNPKNITIDIKRLIGLQVNDPTVQRDILKILPTIYVNNKNGKPYEVATV-KGEKKEPTPEE 165
 DB 127 ISAMVLTKEKTEATLGGKRVTHAVYTPATFNDARQATADACTIAGLVNRIINEPTA 186
 DB 166 VSCNLTGKMKQIAEDYLGKRVTHAVYTPATFNDARQATADACTIAGLVNRIINEPTA 225
 QY 187 AATAYGLDKREGEKNILVLDLGGCTFDVSLITIDNGVEVATNGDTHLGGEDFDORVME 246
 DB 226 AATAYGLDKTEDEHIIIVYDLGGCTFDVSLIENGVEVATNGDTHLGGEDFDVLR 285
 QY 247 HFIKLYKKKTGKQVKNDAVOKLREVEKAKKALSSOHAKRITSEFGEDEPSETLVR 306
 DB 286 HFMQLPQRKIDLVYKNDKAMKLRREKAKKRSLSOTSTRLEIDSEFPGIDSEPTLR 345
 QY 307 AKFEELNMDLFRSTMKPVOKVLEDSDLKSDIDELIVLGGSTRIPIQOOLVKEFFNGKEP 366
 DB 346 AKFEELNMLAFKRTKLPKPEKVLKDSGLQKEDIDIVLGGSTRIPIKVOOLLEKFFNGKA 405
 QY 367 SRGINDPEAVYGAOAVVLSGDDDTDIIVLCYLTIGETVGVGMKRLIPRNVVP 426
 DB 406 SKGINDPEAVYGAOAVVLSGDEGEVYLLDVNALLLTIGETVGVMPPLIKRNPATP 465
 QY 427 TKRSQIFSTASDNPVTITKVEGERPLTKNHLGLTFDLGIPPARGVQOLEVPEID 486
 DB 466 TKRSQIFSTAVDNOKAVRIQVEGERAMVKNLGNFELSDIRAPRGVQOLEVPTALD 525
 QY 487 VNGILRYTADKGTGKKNKTIITNDQNRILTPETIRMYNDAEKRAEDDKLKEKIDITRNE 546

DB 526 ANGLIVTSATDKDTGKSESTITLANDGRISODDIDRMYEVEKKAADAKPKAKSEARNP 565
 QY 547 LESYAVSLKNOQIDKEKRLGKLSDEDEKTEMEAEKEKTEMLESHQD-ADIEDFAKKREL 605
 DB 586 FENFVHYKNSVNG--ELAEIMDEDDKETVLNVNVESELMEDNSDVAAEDPEKKASVF 643
 QY 606 EETVQPIISKLYGSAG 621
 DB 644 KESEVETLARSASQG 659

RESULT 7
 US-08-797-358B-3
 Sequence 3, Application US/08797358B
 Patent No. 6268478
 GENERAL INFORMATION:
 APPLICANT: Adams, John
 TITLE OF INVENTION: INTRACELLULAR VITAMIN D BINDING PROTEIN
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell & Flores LLP
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: United States
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/797,358B
 FILING DATE: 11-Feb-1997
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/011,491
 FILING DATE: 12-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-CE 3165
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 643 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-08-797-358B-3

Query Match 59.2%; Score 1938; DB 4; Length 643;
 Best Local Similarity 61.7%; Pred. No. 1e-147;
 Matches 379; Conservative 107; Mismatches 122; Indels 6; Gaps 5;

QY 13 VGIDLTGTTSCVGVFNKGRVEILANDQGRITPSVATPGEERLIGDAKNQTSNPN 72
 DB 9 VGIDLTGTTSCVGVFOGGRVEILANDQGRITPIVATF-DTNRLVGDAAINQAAALPLN 67
 QY 73 TVFPAKRLIGRTWNDSPVOODIKFLPFKYVEKRTPTIYQVIGGQGTTFAPAEISAVL 132
 DB 68 TVFPAKRLIGRTWNDSPVOODIKFLPFKYVEKRTPTIYQVIGGQGTTFAPAEISAVL 136
 QY 133 TKMKEATAVYLGKRVTHAVYTPATFNDARQATADACTIAGLVNRIINEPTAAATAYG 192
 DB 127 SKMKEATAVYLGQVTHAVYTPATFNDARQATADACTIAGLVNRIINEPTAAATAYG 186

```

OY 193 LDKR-EGEKIILVFDLGGTPEVSLITIDNGYFEVYVATNGTILHGDEDFQORVMEHFIL 251
DB 187 LDRGAGERVLLIFDLGGTPEVSLIDAGYFEVKATVAGDTILHGDEDFQORVMEHFIL 246
OY 252 YKRTGKDYKKNRAVQKLRREVEKAKRALSQHOARIESFEDEGFSETLIRAKFEE 311
DB 247 FRRKHRRKLSMNRRLRLRATCERAKRRTLSSTQATLEIDSLEPGVDFTSTIRAFEE 306
OY 312 LNMDDPRSTKPRVQKVLSDSKSDIDIEIVYGGSTRIPIQOVLVMEFRNGKPEPSGIN 371
DB 307 LCSDFPSTLEPEKGLRQAKLDKAKIHVVYGGSTRIPIVQKLLDDPFGEKELNSIN 366
OY 372 PDEAVAYGAAGVAVLSDG-ODTGDVLVLDVCPPLTIGLIEYGVMTKLIPRNVPTPK 429
DB 367 PDEAVAYGAAGVAVLSDGCKEVRDLDLDVAPLSLGLTAGCGVMTLLDORNAITPTKQ 426
OY 430 SGIPTASDQPTVTIKVEGERPLTNDNLGTFDLTGIPPARGVQIETFEIDVNG 489
DB 427 TQFTTSDQPGVFQYVEGERAMTKONMLGRELSCIPPARGVQIETFEIDVNG 486
OY 490 ILKVLDEKCTGKKNKTTTNDONRLTPELEIRVYNDKFAEDKKLKERIDTRNELS 549
DB 487 ILSTVADRSTGKANKITTTNDKGRLSKEEVRVRAEAOYKADDEAORVAKNSLET 546
OY 550 YVSLKNOIDKCKLGSSEDEKEMKAVEKIEMLSHQDADIEDFKAKKLEIEV 609
DB 547 HVHVKQSL-QEESLADKIPKEDRKHVQDKQGVLANLEIHNQADKEEYHOKRELEQIC 605
OY 610 QPILSKYGSAGPP 623
DB 606 RPIPSRLYGGGVP 619

```

```

; ORIGINAL SOURCE:
; ORGANISM: Murine
; IMMEDIATE SOURCE:
; CLONE: mortalin
; US-08-214-583-2

Query Match 47.38; Score 1547.5; DB 1; Length 679;
Best Local Similarity 51.28; Pred. No. 2.9e-116;
Matches 333; Conservative 101; Mismatches 163; Indels 53; Gaps 14;

OY 10 GTTVGIDLGTTCVGVFKNGRVELIANDQGRITPSVYALTFEGEERLIGDAARNQSLTN 69
DB 53 GAVVIGIDLGTTCVGVFKNGRVELIANDQGRITPSVYALTFEGEERLIGDAARNQSLTN 112
OY 70 PENTVPFARKLIGRTWMDPSVODIFLFPKVEKRTKPYIOWDIGSGGOTKTPAPEISA 129
DB 113 PNTFVATKRIIGRRDDPEVQDKTNVFKIV-RASNDAAVEAHG---KLSPSQIGA 168
OY 130 MYLTKMETAEVILGKVTYHVVTPAYTNDQROATDAGTACLANMRLINEPTAAI 189
DB 169 FVLKMKETAEVILGKVTYHVVTPAYTNDQROATDAGTACLANMRLINEPTAAI 228
OY 190 AYGLDKREGEKNIILVFDLGGTPEVSLITIDNGYFEVYVATNGTILHGDEDFQORVMEHF 249
DB 229 AYGLDKRE-KVIAVVDLGGTPEVSLITIDNGYFEVYVATNGTILHGDEDFQORVMEHF 287
OY 250 KLYKRTGKDYKKNRAVQKLRREVEKAKRALSQHOARIESFEDEGFSETLIRAKFEE 303
DB 288 KEFRRETVGVDLTKDNMALQVRAEAKRACELSSVQTDINLP-VLMDASGRKHLNMK 345
OY 304 LTRAKFELNMDDPRSTKPRVQKVLSDSKSDIDIEIVYGGSTRIPIQOVLVMEFRNGKPEPSGIN 363
DB 346 LTRAKFELNMDDPRSTKPRVQKVLSDSKSDIDIEIVYGGSTRIPIQOVLVMEFRNGKPEPSGIN 404
OY 364 KESPRGGINPDEAVAYGAAGVAVLSDGDOPTGDLVLDVCPPLTIGLIEYGVMTKLIPRNVPTPK 423
DB 405 KESPRGGINPDEAVAYGAAGVAVLSDGDOPTGDLVLDVCPPLTIGLIEYGVMTKLIPRNVPTPK 462
OY 424 VPTKSOIFSTASDQPTVTIKVEGERPLTNDNLGTFDLTGIPPARGVQIETFEIDVNG 483
DB 463 TTPKKSQVSTADQPTVQYVEGERAMTKONMLGRELSCIPPARGVQIETFEIDVNG 522
OY 484 EIDVNGILRYAEDKGTGKKNKTTTNDONRLTPELEIRVYNDKFAEDKKLKERIDTRNELS 541
DB 523 DIDANGIYHSAKDKGTGREGQOIVT-QSSGGLSKDIDEMVYNAKFAEDKKRKEVEA 581
OY 542 -----DTRNELSY-----AYSLKNOIDKCKLGSSEDEKEMKAVEKI 584
DB 582 VDMAEGIIHDTETMEEFKQDLPADDECNLKEIISVRLALLARKDSETEGENIROAA----- 637
OY 585 EMLSHODADIEDFKAKKLEIEVQIPIKSLYGSAGPPNGE--EDTAE 632
DB 638 ---SSLDQASLKLFEVAYKKM-----ASBRGGS-GSSGTGEOKEDQKE 676

RESULT 9
US-08-441-139-4
; Sequence 4, Application US/08441139
; Patent No. 573245
; GENERAL INFORMATION:
; APPLICANT: Wiltup, Dr. Karl D.
; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530

```



```

OY 251 LYKKTGKDVKRDKNRAVOKLRREVAKRALSQHOARLE---IESFEVGE-EDFSETLTR 306
Db 240 SFOKETINLRNDRPMAYOKRKALEKAKIETLRLETDTITLPLFISDSGTAKHLSKLRSR 299
OY 307 AKPEELMDELFRSTMKPVOKVLEDSLK-KSDIDEIVLVGSGSTRIPKIQOLVKEFNGKE 365
Db 300 AKFGLDELIEITIERCKKALSDAGIKDNSKYDEVVLVGGTWRPKYIQHVKNDF-GKE 358
OY 366 PSRGINDEVAAYGAOVAGVLSGDDDTGDLVLDVCPPLTGIEFVGVMMKILPRNTVV 425
Db 359 PCQGVNDEVAAYGAALOGGILTG--DVARDVLLDVAPLSLGIEITLGGVFPPLERTTTI 416
OY 426 PTKSQSTFASDNQPTVYIKYEGEERPLTKDNHLGTFDLTGIPAPRGVPOIEVTFEI 485
Db 417 PTKSQVSTAEAGQAVYIKYQGERKMAIDNKLGGFSPHAPRGVPOIEVTFDI 476
OY 486 DVNGLLVTAEKGTGKNNKITTTNDQNLTPPEELERVYNDAEKFAEDKKLRIDTRN 545
Db 477 DANGIVHSAKDKASGEQIRK-QSSGGLSDEPIKKNVKAQADRAEDDEKRAKHVELKN 535
Db 546 ELESYAVSLKNQIGDKELGKLSSEDKETMEKAVEEKIEML 587
Db 536 SSEGILHVSERKSLKD--YGDYKAGADKSNIESAIKIDRECL 574

```

```

RESULT 11
US-08-472-534-5
: Sequence 5, Application US/08472534
: Patent No. 5918620
: GENERAL INFORMATION:
: APPLICANT: Hamel, Josee
: APPLICANT: Brodeur, Bernard R
: APPLICANT: Martell, Denis
: TITLE OF INVENTION: HEAT SHOCK PROTEIN HSP72 FROM
: TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Neave
: STREET: 1251 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10020
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/472,534
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Haley Jr, James F
: REGISTRATION NUMBER: 27,794
: REFERENCE/DOCKET NUMBER: Biovac-2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-596-9000
: TELEFAX: 212-596-9090
: TELEX: 14-8367
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 607 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-472-534-5

```

```

Query Match 44.18; Score 1444; DB 2; Length 607;
Best Local Similarity 49.68; Pred. No. 5.2e-108;
Matches 308; Conservative 99; Mismatches 170; Indels 44; Gaps 11;

```

```

OY 12 VVGIDLGITTCVGVCKNGRVELIANDQGNRIPTSPVATPPEGEHLIDDAKNOITSNPE 71
Db 4 IIGIDLGITNSAAVAVLGTESKIIANPEGKRTTPSVVS-KKGEIIVDAKRAQAVNPD 62
OY 72 NTVFDARLIGRTWNPDSVOODIKRPLPFVYERKTKPYIQVDIGGGQRTFAPEISAMV 131
Db 63 -TVISIKSMQTS-----EK-----VANG---KEYTPOEISMI 93
OY 132 LYKKEAEVAVLGKVVTHAVTVPAYFNDAQRAATGACTIGLVMBITNEPPAAIAY 191
Db 94 LQYLGKAEIDYLGKTKAVITVPAFNDAQRAATGACTIGLVMBITNEPPAAIAY 153
OY 192 GLDRREGKNIIVFDLGGTFDVSLLTIDNGVEVAVTNGDTHLGEDEFDQRYMEHFJKL 251
Db 154 GLDRTRKEEKIVFDLGGTFDVSILLEDGVEYDLSTAGDNKLGGDDFDKIIDHVAE 213
OY 252 LYKKTGKDVKRDKNRAVOKLRREVAKRALSQHOARLEIESFEGE----DFSETLTR 307
Db 214 FKKEGIDSLDRAKAMORLDAAEKAKKDLGVTSTSLPRTTAGEAGPLHLENTLRA 273
OY 308 KPEELMDELFRSTMKPVOKVLEDSLK-KSDIDEIVLVGSGSTRIPKIQOLVKEFNGKEPS 367
Db 274 KFDDLTRDLVETKVPVROALSDAGLSLSEIDVILVGGSTRIPAVEAVKA-ETGKEPN 332
OY 368 RGINPDEVAAYGAOVAGVLSGDDDTGDLVLDVCPPLTGIEFVGVMMKILPRNTVVPT 427
Db 333 KSNPDEVAAYGAALOGGILTG--DVKDVLVLDVPLSLGIEITLGGVFPPLERTTTIPT 390
OY 428 KKSQIFSTASDNQPTVYIKYEGEERPLTKDNHLGTFDLTGIPAPRGVPOIEVTFEIDV 487
Db 391 SKSQVSTAEADNCPAVDIIHVQGERPMADNKTITDFHFLTDIPAPRGVPOIEVTFDIDK 450
OY 488 NGILRVTAEDKGTGKNNKITTTNDQNLTPPEELERVYNDAEKFAEDKKLRIDTRNEL 547
Db 451 NGIYSVAKADLTOKREOTIYIQNSG-LTDEBIDMKADAEANASDKKREEDVLNREV 509
OY 548 ESYAVSLKNQIGDKELGKLSSEDKETMEKAVEEKIEMLSHODADIEDPKAKKELEE 607
Db 510 DOAIFATEKTIKETE--GKGPADERDAQAALD--DLKKAQEDNNIDMKAKLEALNE 563
OY 608 IVOPILSKLYGSAGPPTGEE 628
Db 564 KAQGLAVKLYEQAAAAQAQAE 584

```

```

RESULT 12
5240706-1
: Patent No. 5240706
: APPLICANT: FAULDS, DARYL
: TITLE OF INVENTION: INTRANASSAL ADMINISTRATION OF MYOGLIPLASMA
: HYPONEUMONIAE ANTIGEN
: NUMBER OF SEQUENCES: 22
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/334,586
: FILING DATE: 07-APR-1989
: SEQ ID NO:1:
: LENGTH: 600
: 5240706-1

```

```

Query Match 39.88; Score 1303.5; DB 6; Length 600;
Best Local Similarity 45.38; Pred. No. 1e-96;
Matches 282; Conservative 106; Mismatches 187; Indels 47; Gaps 11;
OY 12 VVGIDLGITTCVGVCKNGRVELIANDQGNRIPTSPVATPPEGEHLIDDAKNOITSNPE 71
Db 6 IIGIDLGITNSAAVAVLGTESKIIANPEGKRTTPSVVS-KKGEIIVDAKRAQAVNPD 62
OY 72 NTVFDARLIGRTWNPDSVOODIKRPLPFVYERKTKPYIQVDIGGGQRTFAPEISAMV 131
Db 65 -AIASIKRLMGTD-----KTVANERDYI-----PEISAKI 95

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 339 amino acids
Type: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5958727e
US-08-928-692-52

Query Match      34.8%; Score 1140.5; DB 2; Length 339;
Best Local Similarity 67.8%; Pred. No. 5.8e-84;
Matches 229; Conservative 44; Mismatches 62; Indels 3; Gaps 3;

QY 13 VGIDLGTVYSCVGFKNCRVREIIANDGNGRTPTSPYVAFTEPGERLLIGDAKNOIGNSPEN 72
Db 5 VGIDLGTVYSCVAFHFNDRVIDIANDGNGRTPTSPVAFTEPGERLLIGDAKNOIGNSPEN 63

QY 73 TVFDARKLIGRTNDSVQODIKFLPFVYVKKTKRPYIOVDIGSGOTKTFAPETISAWL 132
Db 64 TVFDARKLIRGNNDSEVQADKHPFLIIVDVKPQIOYEF-KGELTKNFTPEQISSWL 122

QY 133 TKKKEFLAEVYLKKKTHAVYVPATFNDAQSQATKDKAGTAGLVNMRITNEPFAAIAVG 192
Db 123 GKKEKEPAESTLGAKVADVAVTPAFNDSORQATKAGTAGLVNMRITNEPFAAIAVG 182

QY 193 LDRKEGKRLIVFDLGSGTFEDVSLTTINGVFEEVATNNGDTHLGGEDFDQRVMEHFRLKY 252
Db 183 LDKKGEKEHFLFDLGSGTFEDVSLTFIEDGAFEVKATAGDTHLGGEDFDQRLVNHFIQEF 242

QY 253 KKKTKGADVRRDNRNAVOKLREVEKAKKRALSSOHQARIEFIESFEGEDFSETLTRAFEEFL 312
Db 243 KKKKKKOLSTNOALRRRLTRFACESQENFVSSA-OTVSIEDSKNEGIDFTTSTIRAFEEFL 301

QY 313 NMDLFRSTMKPVOKVLEDSDLKKSIDIEIVLVGSGSTR1 350
Db 302 CADLFRSTLDPEVEKVLRLDAKLDKSGVDIEIVLVGSGSTR1 339

RESULT 14
5196523-13
Patent No. 5196523
APPLICANT: LEE, AMY S.
TITLE OF INVENTION: CONTROL OF GENE EXPRESSION BY GLUCOSE,
CALCIUM AND TEMPERATURE
NUMBER OF SEQUENCES: 28
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-07/354,988
FILING DATE: 19-MAY-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 282,880
FILING DATE: 05-DEC-1988
APPLICATION NUMBER: 690,951
FILING DATE: 01-JAN-1985
SEQ ID NO:13:
LENGTH: 187
5196523-13

Query Match      28.8%; Score 941; DB 6; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.7e-68;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 450 GERLTLNDNHLTLFTDLTGIPAPAPRGVQIIEVTFIDVNGCLTRKLTADKGTGKNNITTT 509
Db 1 GERLTLNDNHLTLFTDLTGIPAPAPRGVQIIEVTFEIDVNGCLTRKLTADKGTGKNNITTT 60

QY 510 NDONRLTPEELIERMVDAAKFAEDDKLKERIDTRNELESYAASLKNQIDDKERLGGKLS 569
Db 61 NDONRLTPEELIERMVDAAKFAEDDKLKERIDTRNELESYAASLKNQIDDKERLGGKLS 120

QY 570 SEDKETMEKAVEKEIKWLESODADIEDFKAKKKELEIYVPIITSKIYSGAGPPPTGEED 629
Db 121 SEDKETMEKAVEKEIKWLESODADIEDFKAKKKELEIYVPIITSKIYSGAGPPPTGEED 180

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:13:09 ; Search time 34.99 Seconds

(without alignments)
1754.818 Million cell updates/sec

Title: US-09-806-955A-1

Perfect score: 3273
Sequence: 1 MEEDKEDVGTWVGIDLTGTTSCVGVFKNGRVEIIANDOGNRIRTPSYVAFTPEGERLIGDA 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

1 number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database :

PIR-71:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	98.0	654	2 A27414	dnak-type molecule
2	3201	97.8	655	2 A37048	dnak-type molecule
3	3200	97.8	654	1 HHRGB	dnak-type molecule
4	3182.5	97.2	653	2 A29821	dnak-type molecule
5	3168	96.8	652	2 T50342	dnak-type molecule
6	2722.5	83.2	667	2 S24782	dnak-type molecule
7	2719.5	83.1	667	2 D44261	dnak-type molecule
8	2710	82.8	656	2 JN0666	dnak-type molecule
9	2618.5	80.0	661	2 T15513	heat shock 70K pro
10	2594.5	79.3	651	2 A32475	dnak-type molecule
11	2559	78.2	657	2 T34037	heat shock 70K pro
12	2356.5	72.0	667	2 S21879	dnak-type molecule
13	2347.5	71.7	666	2 S21880	dnak-type molecule
14	2332.5	71.3	669	2 T04080	dnak-type molecule
15	2330	71.0	668	2 T04078	dnak-type molecule
16	2322.5	70.4	663	2 T03581	dnak-type molecule
17	2305.5	70.2	668	2 S71171	dnak-type molecule
18	2298	69.8	672	2 T43723	dnak-type molecule
19	2285	69.6	672	2 T43716	dnak-type molecule
20	2278	69.1	664	2 T45298	dnak-type molecule
21	2261	68.5	664	2 T06598	dnak-type molecule
22	2240.5	68.3	655	2 H86222	hypothetical prote
23	2237	68.2	658	2 S38690	glucose-regulated
24	2231	68.1	662	2 T50464	78 kd glucose regu
25	2227.5	67.7	663	2 T38155	dnak-type molecule
26	2216	67.7	668	2 T05741	dnak-type molecule
27	2214.5	67.4	666	2 T06358	dnak-type molecule
28	2205	66.4	682	1 H8BYK2	dnak-type molecule
29	2174				

30	2156	65.9	663	2 S20877	dnak-type molecule
31	2108	64.4	651	2 JC7132	heat shock protein
32	2100	64.2	655	2 A48468	dnak-type molecule
33	2083	63.6	646	2 S07197	dnak-type molecule
34	2083	63.6	646	2 JC4853	dnak-type molecule
35	2083	63.6	646	2 A27027	dnak-type molecule
36	2080	63.6	641	2 PC7036	heat shock protein
37	2079	63.5	647	2 T41121	heat shock protein
38	2079	63.5	650	2 S11456	dnak-type molecule
39	2078	63.5	645	2 S41372	dnak-type molecule
40	2076.5	63.4	646	2 A45935	dnak-type molecule
41	2076.5	63.4	640	2 S37394	dnak-type molecule
42	2074.5	63.4	648	2 S53498	dnak-type molecule
43	2074.5	63.4	656	2 S51712	dnak-type molecule
44	2074	63.4	646	2 A35922	dnak-type molecule
45	2072	63.3	639	2 S20139	dnak-type molecule

ALIGNMENTS

RESULT 1
A27414
dnak-type molecular chaperone GRP78 precursor - Chinese hamster
N:Alternate names: glucose-regulated 78K protein; Ig heavy chain-binding protein
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 30-Jun-1988 *sequence revision 30-Jun-1988 *text-change 20-Aug-1999
C:Accession: A27414
R:Ring, J.; Woodem, S.K.; Ritz, R.; Kelleher, K.; Kaufman, R.J.; Lee, A.S.
Gene 55, 147-152, 1987
A:Title: The nucleotide sequence encoding the hamster 78-kDa glucose-regulated protel
A:Reference number: A27414; MUID:87305586
A:Accession: A27414
A:Molecule type: mRNA
A:Residues: 1-654 <TIN>
A:Cross-references: GB:M17169; NID:g191090; PIDN:AAA51448.1; PID:g304510
C:Genetics:
A:Gene: GRP78
C:Function:
A:Description: involved in protein folding and assembling/disassembling of protein co
A:Superfamily: heat shock protein 70
C:Keywords: ATP; heterotetramer; immunoglobulin; molecular chaperone
F:1-16/Domain: signal sequence *status predicted <SIS>
F:19-654/Product: Ig heavy chain binding protein *status predicted <MAT>
F:651-654/Region: endoplasmic reticulum retention signal

Query Match 98.0% Score 3209: DB 2: Length 654:
Best Local Similarity 99.7% Pred. No. 8.1e-155:
Matches 629: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

QY 2 EBDKEDVGTWVGIDLTGTTSCVGVFKNGRVEIIANDOGNRIRTPSYVAFTPEGERLIGDA 61
|||||
DB 20 EBDKEDVGTWVGIDLTGTTSCVGVFKNGRVEIIANDOGNRIRTPSYVAFTPEGERLIGDA 79
|||||

QY 62 ANKQGLSPNENYVPAKRLIGRTMNDSPVOODIKFLPKYVEKKTPYIOVIGGGQRT 121
|||||
DB 80 ANKQGLSPNENYVPAKRLIGRTMNDSPVOODIKFLPKYVEKKTPYIOVIGGGQRT 139
|||||

QY 122 FAPETISAVYLTKMKETAAYLGKKVTHAVVTPAYPNDAROATRDAGTGLANMYRII 181
|||||
DB 140 FAPETISAVYLTKMKETAAYLGKKVTHAVVTPAYPNDAROATRDAGTGLANMYRII 199
|||||

QY 182 NEPTAAALNYGDKREGKNTLVLDGGTFDVSLLTIDNGVEVAATNGDHILGGEEDF 241
|||||
DB 200 NEPTAAALNYGDKREGKNTLVLDGGTFDVSLLTIDNGVEVAATNGDHILGGEEDF 259
|||||

QY 242 QRYMEHFIKLYKKKTGKVDNRDNRAVOKLRREYKAKRALSSOQARIEISFEDEGDS 301
|||||
DB 260 QRYMEHFIKLYKKKTGKVDNRDNRAVOKLRREYKAKRALSSOQARIEISFEDEGDS 319
|||||

QY 302 ETLTRAKFEELNMDLFRSTMKPVOKVLEDSDLKSDIDELVVLGGSTRIPKIQOLVYKEF 361
|||||

```

Db 320 ETLTRAKFPEELNMDLFRSTMKPVQKVLSDLSLKKSDIDEIVLGVGSTRIPKIQOLYKEFF 379
Oy 362 NGKEPSRGINPDEAVAYGAAGVAGVLSGDOPTGDLVLDVCPRLTGIEYGVWTKLIR 421
Db 380 NGKEPSRGINPDEAVAYGAAGVAGVLSGDOPTGDLVLDVCPRLTGIEYGVWTKLIR 439
Oy 422 NTVPPTKKSQIFSTASDNQPTVTIKVYEGEERPLTKDNHLLGTDTLGTIPAPRGVQIEV 481
Db 440 NTVPPTKKSQIFSTASDNQPTVTIKVYEGEERPLTKDNHLLGTDTLGTIPAPRGVQIEV 499
Oy 482 TFEIDVNGILRTVAEDKGTGNKNTITINDQNRLLTPEIERMNVDAEKFAEEDKKLKERI 541
Db 500 TFEIDVNGILRTVAEDKGTGNKNTITINDQNRLLTPEIERMNVDAEKFAEEDKKLKERI 559
Oy 542 DTRNELESAVYSLKNOIGDKELGKLSSEDEKETMEKAVEEIKEMLESQDADIEDFKAK 601
Db 560 DTRNELESAVYSLKNOIGDKELGKLSSEDEKETMEKAVEEIKEMLESQDADIEDFKAK 619
Oy 602 KKELEIVQPIISKLYGSAGPPPTGEEDTAE 632
Db 620 KKELEIVQPIISKLYGSAGPPPTGEEDTAE 650

```

RESULT 2

```

A37048
dnak-type molecular chaperone grp78 precursor - mouse
N:Alternate names: B1p, glucose-regulated 78k protein, Ig heavy chain-binding protein
C:Species: Mus musculus (house mouse)
C>Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 20-Aug-1999
C:Accession: A37048; B37048; J00094; I49137; A31934
R:Kozutsumi, Y.; Normington, K.; Press, E.; Slaughter, C.; Sambrook, J.; Gething, M.J.
J. Cell Sci. Suppl. 11, 115-137, 1989
A:Title: Identification of immunoglobulin heavy chain binding protein as glucose-regulat
A:Reference number: A37048; MUID:90130686
A:Accession: A37048
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-655 <K21>
A:Accession: B37048
A:Molecule type: Protein
A:Residues: 20-41, 'X', 43-46 <K22>
R:Parfett, C.L.J.; Hofbauer, R.; Brudzynski, K.; Edwards, D.R.; Denhardt, D.T.
Gene 82, 291-303, 1989
A:Title: Differential screening of a cDNA library with cDNA probes amplified in a heteroc
A:Reference number: J00094; MUID:90060818
A:Accession: J00094
A:Molecule type: mRNA
A:Residues: 488-590, 'G', 592-595, 'E', 597-655 <PAR>
A:Cross-references: GB:M30779; NID:g193644; PIDN:AAA37742.1; PID:g193645
R:Blum, J.B.; Mote, P.L.; Walford, R.L.; Splindler, S.R.
Gene 158, 225-229, 1995
A:Title: Structure and regulation of the mouse GRP78 (B1p) promoter by glucose and calci
A:Reference number: I49137; MUID:95331621
A:Accession: I49137
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-28 <REBS>
A:Cross-references: EMBL:U16277; NID:g829364; PIDN:AAA37742.1; PID:g829365
R:Haas, I.G.; Moe, T.
Proc. Natl. Acad. Sci. U.S.A. 85, 2250-2254, 1988
A:Title: cDNA cloning of the immunoglobulin heavy chain binding protein.
A:Reference number: A31934; MUID:88176922
A:Accession: A31934
A:Molecule type: mRNA
A:Residues: 514-589, 'D', 591-595, 'E', 597-655 <HAA>
A:Cross-references: GB:M19351; NID:g192231; PIDN:AAA37315.1; PID:g387113
C:Genetics:
A:Gene: grp78
C:Function:
A:Description: Involved in protein folding and assembling/disassembling of protein comp
C:Superfamily: heat shock protein 70
C:Keywords: ATP; endoplasmic reticulum; glycoprotein; heterotetramer; molecular chaperon
F:1-Ig/Domain: signal sequence #status predicted <SIG>

```

F:20-655/Product: dnak-type molecular chaperone grp78 #status experimental <MAT>
F:652-655/Region: endoplasmic reticulum retention signal

Query Match 97.8%; Score 3201; DB 2; Length 655;

Best Local Similarity 99.4%; Pred. No. 2, 1e-154;

Matches 627; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

Oy 2 EEKKKEDVGTVGIDIGTTCVGVFKNGRVELLIANDQGNRTGPSYAFPPPEERLIGDA 61
Db 21 EEKKKEDVGTVGIDIGTTCVGVFKNGRVELLIANDQGNRTGPSYAFPPPEERLIGDA 80
Oy 62 AKNOLTSNPENTVFDKAKRLIGFTWMDPSVQDDIKFLPEKVEKKTKPYIOVDIGGQTKT 121
Db 81 AKNOLTSNPENTVFDKAKRLIGFTWMDPSVQDDIKFLPEKVEKKTKPYIOVDIGGQTKT 140
Oy 122 FAREEISAMVTRKMKETAAYLGKKVTAAVTVPAVFMDAOROKTKAGTITAGLNMWRT 181
Db 141 FAREEISAMVTRKMKETAAYLGKKVTAAVTVPAVFMDAOROKTKAGTITAGLNMWRT 200
Oy 182 NEPTAAAIAYGLDKRGEKNILVPDLGGTFDVSLLTIDNCGFEVATNGDTILGDEPD 241
Db 201 NEPTAAAIAYGLDKRGEKNILVPDLGGTFDVSLLTIDNCGFEVATNGDTILGDEPD 260
Oy 242 QRYMEHPIKLYKKKTKGVKDNRAVOKLREVEKAKRALSSOQARTEIESFVEGEDPS 301
Db 261 QRYMEHPIKLYKKKTKGVKDNRAVOKLREVEKAKRALSSOQARTEIESFVEGEDPS 320
Oy 302 ETLTRAKFPEELNMDLFRSTMKPVQKVLSDLSLKKSDIDEIVLGVGSTRIPKIQOLYKEFF 361
Db 321 ETLTRAKFPEELNMDLFRSTMKPVQKVLSDLSLKKSDIDEIVLGVGSTRIPKIQOLYKEFF 380
Oy 362 NGKEPSRGINPDEAVAYGAAGVAGVLSGDOPTGDLVLDVCPRLTGIEYGVWTKLIR 421
Db 381 NGKEPSRGINPDEAVAYGAAGVAGVLSGDOPTGDLVLDVCPRLTGIEYGVWTKLIR 440
Oy 422 NTVPPTKKSQIFSTASDNQPTVTIKVYEGEERPLTKDNHLLGTDTLGTIPAPRGVQIEV 481
Db 441 NTVPPTKKSQIFSTASDNQPTVTIKVYEGEERPLTKDNHLLGTDTLGTIPAPRGVQIEV 500
Oy 482 TFEIDVNGILRTVAEDKGTGNKNTITINDQNRLLTPEIERMNVDAEKFAEEDKKLKERI 541
Db 501 TFEIDVNGILRTVAEDKGTGNKNTITINDQNRLLTPEIERMNVDAEKFAEEDKKLKERI 560
Oy 542 DTRNELESAVYSLKNOIGDKELGKLSSEDEKETMEKAVEEIKEMLESQDADIEDFKAK 601
Db 561 DTRNELESAVYSLKNOIGDKELGKLSSEDEKETMEKAVEEIKEMLESQDADIEDFKAK 620
Oy 602 KKELEIVQPIISKLYGSAGPPPTGEEDTAE 632
Db 621 KKELEIVQPIISKLYGSAGPPPTGEEDTAE 651

```

RESULT 3

```

HHR7GB
dnak-type molecular chaperone precursor - rat
N:Alternate names: B1p, glucose-regulated 78k protein, Ig heavy chain-binding protein
N:Contains: steroidogenesis-activator polypeptide
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 22-Jun-1999
A:Accession: A23948; A60134; A26257
R:Munro, S.; Pelham, H.R.B.
Cell 46, 291-300, 1986
A:Title: An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated p
A:Reference number: A23948; MUID:86245075
A:Accession: A23948
A:Molecule type: mRNA
A:Residues: 1-654 <MUN>
A:Cross-references: GB:M14050; NID:g203150; PIDN:AAA40817.1; PID:g203151
R:Pedersen, R.C.; Brownlee, A.C.
Science 226, 188-190, 1987
A:Title: Steroidogenesis-activator polypeptide isolated from a rat Leydig cell tumor.
A:Reference number: A60134; MUID:87177981

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:17:49 ; Search time 19.71 Seconds
(without alignment)
1255.291 Million cell updates/sec

Title: US-09-806-955A-1

Perfect score: 3273

Sequence: 1 MEEDKEDGVGVGIDLCCT.....AGPPPEEDTALHHHHH 639

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

1 number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3216	98.3	654	1 GR78_HUMAN	P11021 homo sapien
2	3209	98.0	654	1 GR78_MESAU	P07823 mesocricetu
3	3205	97.9	655	1 GR78_MOUSE	P20029 mus musculu
4	3200	97.8	654	1 GR78_RAT	P06761 ratu muscu
5	3168	96.8	652	1 GR78_CHICK	O90593 gallus gall
6	3069	93.8	658	1 GR78_XENLA	O91883 xenopus lae
7	2722.5	83.2	667	1 GR78_APLCA	O16956 aplysia cal
8	2708	82.7	656	1 HS7C_DROME	P28844 drosophila
9	2594.5	79.3	661	1 HS7C_CAERL	P27420 caenorhabdi
10	2356.5	72.0	667	1 BIP5_TOBAC	O03684 nicotiana t
11	2347.5	71.7	668	1 BIP5_TOBAC	O03685 nicotiana t
12	2334.5	71.3	663	1 BIP3_MAIZE	O24581 zea mays (m
13	2328.5	71.1	666	1 BIP_LYCES	O44918 lycopersico
14	2328.5	71.1	668	1 BIP_SPIOL	O42434 spincia ol
15	2322.5	71.0	663	1 BIP3_MAIZE	P24067 zea mays (m
16	2301	70.3	669	1 BIP1_ARATH	O91843 arabidopsis
17	2299	70.2	668	1 BIP2_ARATH	P78695 arabidopsis
18	2227.5	68.1	663	1 GR78_NEUCR	P78695 arabidopsis
19	2216	67.7	663	1 GR78_SCHPO	P36604 schizosacch
20	2195	67.1	670	1 GR78_YARLI	O99170 yarowia li
21	2174	66.4	682	1 GR78_YEAST	P14474 saccharomyc
22	2100	64.2	655	1 GR78_PLAFO	O08866 plasmodium
23	2085.5	63.7	641	1 HS71_YEAST	P10591 saccharomyc
24	2083	63.6	646	1 HS7C_HUMAN	P11142 homo sapien
25	2083	63.6	646	1 HS7C_MOUSE	P08109 mus musculu
26	2079	63.5	650	1 HS7C_BOVIN	P19120 bos taurus
27	2078	63.5	644	1 HS71_PICAN	P53421 picchia angu
28	2076.5	63.4	640	1 HS7C_DICDI	P36415 dictyosteli
29	2074.5	63.4	655	1 HS71_CANAL	P41797 candida alb
30	2074	63.4	646	1 HS7C_CRIGR	P19378 cricetus
31	2072	63.3	638	1 HS72_YEAST	P19592 saccharomyc
32	2065.5	63.1	679	1 GR78_KLULA	P22010 kluyveromyc
33	2065	63.1	648	1 HS71_PUCGR	O01877 puccinia gr

34	2064	63.1	641	1 HS74_YEAST	P22202 saccharomyc
35	2063.5	63.0	649	1 HS73_YEAST	P09435 saccharomyc
36	2060.5	63.0	646	1 HS70_NEUCR	O01233 neurospora
37	2060	62.9	643	1 HS71_SCHPO	O10265 schizosacch
38	2059	62.9	645	1 HS70_SOYBN	P26413 glycine max
39	2053.5	62.7	651	1 HS7C_PENNY	P09189 petunia hyb
40	2052	62.7	652	1 HS7D_MANS	O94639 manduca sex
41	2049	62.6	651	1 HS70_ONCHR	P08108 oncomyrmehu
42	2046	62.5	641	1 HS72_BOVIN	O27965 bos taurus
43	2045	62.5	651	1 HS71_ARATH	P22953 arabidopsis
44	2043.5	62.4	639	1 HS72_HUMAN	P54652 homo sapien
45	2039.5	62.3	641	1 HS71_HUMAN	P08107 homo sapien

ALIGNMENTS

RESULT	ID	GR78_HUMAN	STANDARD:	PRT:	654 AA.
AC	P11021	09NPFL:			
DT	01-JUL-1989	(Rel. 11, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	78 kDa glucose-regulated protein precursor (GRP 78) (Immunoglobulin heavy chain binding protein) (BiP) (Endoplasmic reticulum lumenal Ca2+ binding protein grp78).				
DE	HSP45 OR GRP78.				
GN	Homo sapiens (Human).				
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]	SEQUENCE FROM N.A.			
RP	MEDLINE=8628347; PubMed=2840249;				
RX	Ting J., Lee A.S.;				
RA	Human gene encoding the 78,000-dalton glucose-regulated protein and its pseudogene: structure, conservation, and regulation.*;				
RT	DNA 7:275-286(1988).				
RL	[2]	SEQUENCE FROM N.A.			
RN	[3]	SEQUENCE FROM N.A.			
RP	TISSUE=Fibroblast;				
RC	Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.;				
RA	GRP78 is involved in the quality control of the IgL-receptor.*;				
RT	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.				
RL	[4]	SEQUENCE FROM N.A.			
RN	Bernandez-Pajardo A., Ijewellyn D.H., Campbell A.K., Errington R.R.;				
RA	*Sequence differences between human grp78/BiP isolated from HeLa cells and previously reported human sequences.*;				
RT	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.				
RL	[5]	SEQUENCE OF 1-25 FROM N.A.			
RN	MEDLINE=93126087; PubMed=1480470;				
RA	Chao C.C.K., Lin-Chao S.;				
RT	*A direct-repeat sequence of the human BiP gene is required for A23187-mediated inducibility and an inducible nuclear factor binding.*;				
RL	Nucleic Acids Res. 20:6481-6485(1992).				
RN	[6]	SEQUENCE OF 22-38.			
RP	TISSUE=Breast carcinoma;				
RC	MEDLINE=97295304; PubMed=9150946;				
RA	Rasmussen R.K., Ji H., Eddes J.S., Moritz R.L., Reid G.E.;				
RT	*Two-dimensional electrophoretic analysis of human breast carcinoma proteins: mapping of proteins that bind to the SH3 domain of mixed lineage kinase MLK2.*;				
RL					

```

RL Electrophoresis 18:588-598(1997).
RN [7]
RP SEQUENCE OF 19-40.
RC TISSUE=Colon carcinoma;
RX MEDLINE=97295306; PubMed=9150948;
RT J1 H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
RT "A two-dimensional gel database of human colon carcinoma proteins.";
RL Electrophoresis 18:605-613(1997).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL: M19645; AAA52614.1; -
EMBL: X87949; CAA61201.1; -
EMBL: AJ271729; CAB71335.1; -
EMBL: AF216292; AAF42836.1; -
EMBL: X59669; CAA42595.1; -
PIR: A29821; A29821.
HSSP: P19120; 3HSC.
SWISS-2DPAGE: P11021; HUMAN.
PMMA-2DPAGE: P11021; -
PNCI-2DPAGE: P11021; -
Slerna-2DPAGE: P11021; -
MIM: I38120; -
InterPro: IPR000886; ER_target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PR00012; HSP70.1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET.1.
DR PROSITE: PS00297; HSP70.1; 1.
DR PROSITE: PS00329; HSP70.2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
KW ATP-binding; Endoplasmic reticulum; Signal.
FT SIGNAL 1 18
FT CHAIN 19 654 78 KDA GLUCOSE-REGULATED PROTEIN.
FT SITE 651 654 PREVENT SECRETION FROM ER.
FT CONFLICT 297 297 MISSING (IN REF. 1 AND 2).
FT CONFLICT 418 418 D -> H (IN REF. 1 AND 2).
FT CONFLICT 439 439 R -> S (IN REF. 1 AND 2).
FT CONFLICT 447 447 K -> N (IN REF. 1 AND 2).
SEQUENCE 654 AA: 72333 MW: 5987085832A00 CRC64;

```

```

Query Match 98.3%; Score 3216; DB 1; Length 654;
Best local similarity 100.0%; Pred. No. 3.2e-152;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 260 QVMEHFILYKKTGDKVKNRAVOKLRREVEKAKAASSQHOARIESEYEGEDFS 319
DB 302 ETLTRAKFELNMDLFRSTMKPVOKVLEDSDLKRSIDIEYLVGSGTRIRKIQOLVKEFF 361
DB 320 ETLTRAKFELNMDLFRSTMKPVOKVLEDSDLKRSIDIEYLVGSGTRIRKIQOLVKEFF 379
DB 362 NGKEPARGINPDEAVYGAAYGAVLSGDQDPTGDLVLDVCPPLTGLIETVGGVMTLIR 421
DB 380 NGKEPARGINPDEAVYGAAYGAVLSGDQDPTGDLVLDVCPPLTGLIETVGGVMTLIR 439
DB 422 NTVPYRKQISFSTASDNPVTIVYEGSERPLTKDNHLGTFDLTGIPAPRGVQIEV 481
DB 440 NTVPYRKQISFSTASDNPVTIVYEGSERPLTKDNHLGTFDLTGIPAPRGVQIEV 499
DB 482 TFEIDVNGILRYAEDKGGKGNKTTITNDQNRLLTPEIERMVNDAKFAEEDKIKLERI 541
DB 500 TFEIDVNGILRYAEDKGGKGNKTTITNDQNRLLTPEIERMVNDAKFAEEDKIKLERI 559
DB 542 DTRNELESTAYSLKNOIGKELGKLSSEDEKTEMEKAVEKIEWLESHODADIEDPKAK 601
DB 560 DTRNELESTAYSLKNOIGKELGKLSSEDEKTEMEKAVEKIEWLESHODADIEDPKAK 619
DB 602 KKELEEIYQPIISKLYSGAPPTGEEDTAE 632
DB 620 KKELEEIYQPIISKLYSGAPPTGEEDTAE 650

```

```

RESULT 2
GR78_MESAU STANDARD: PRT; 654 AA.
ID GR78_MESAU
AC P07823;
DT 01-AUG-1988 (rel. 08, Created)
DT 01-AUG-1988 (rel. 08, Last sequence update)
DT 01-NOV-1997 (rel. 35, Last annotation update)
DE 78 kDa glucose-regulated protein precursor (GRP 78) (Immunoglobulin
DE heavy chain binding protein) (BiP).
OS HSP45 OR GRP78.
GN Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87305586; PubMed=3623104;
RA Ting J., Wooden S.K., Kriz R., Kelleher K., Kaufman R.J., Lee A.S.;
RT "The nucleotide sequence encoding the hamster 78-kDa
RT glucose-regulated protein (GRP78) and its conservation between
RT hamster and rat.";
RL Gene 55:147-152(1987).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL: M17169; AAA51448.1; -
PIR: A27414; A27414.
HSSP: P19120; 3HSC.
DR InterPro: IPR000886; ER_target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PR00012; HSP70.1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET.1.
DR PROSITE: PS00297; HSP70_1; 1.

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:15:59 ; Search time 50.75 Seconds

(without alignments)
2178.202 Million cell updates/sec

Title: US-09-806-955a-1

Sequence: 1 MEDKKEDGTVGIDCTT.....AGPPTEGDDTAALHHHHH 639

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 562222 seqs, 172994929 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PLANT:*
10: SP_PODENT:*
11: SP_VIRUS:*
12: SP_VIRUS:*
13: SP_VIRUS:*
14: SP_VIRUS:*
15: SP_VIRUS:*
16: SP_VIRUS:*
17: SP_VIRUS:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3273	100.0	639	4 Q9UK02	Q9UK02 homo sapien
2	3199	97.7	635	11 Q9DC41	Q9DC41 mus musculu
3	2994	91.5	658	13 Q91688	Q91688 xenopus lae
4	2725	83.3	658	5 Q76180	Q76180 bombyx mori
5	2717	83.0	656	5 Q9VYU3	Q9VYU3 drosophila
6	2607.5	79.7	656	5 Q62568	Q62568 suberites d
7	2559	78.2	657	5 Q02085	Q02085 ctenophora
8	2534	77.4	651	5 Q24798	Q24798 echinococu
9	2525.5	77.2	649	5 Q24895	Q24895 echinococu
10	2460	75.2	648	5 Q45038	Q45038 schistosoma
11	2345.5	71.7	665	10 Q9M4E8	Q9M4E8 cucumis sat
12	2330	71.2	668	10 Q22639	Q22639 glycine max
13	2325.5	70.1	657	5 Q961X6	Q961X6 trichinella
14	2313.5	70.7	675	10 Q40924	Q40924 pseudotsuga
15	2305.5	70.4	663	10 Q24182	Q24182 oryza sativ
16	2292	70.0	668	10 Q9FSY7	Q9FSY7 corytus ave

17	2288.5	69.9	665	10 Q941K4	Q941K4 scherffelia
18	2285	69.8	672	3 Q14453	Q14453 aspergillus
19	2278	69.6	672	3 Q13280	Q13280 aspergillus
20	2273	69.4	668	5 Q9U540	Q9U540 toxoplasma
21	2263	69.1	669	3 Q9C1C1	Q9C1C1 cronartium
22	2261	68.1	642	5 Q61001	Q61001 toxoplasma
23	2246.5	68.6	672	3 Q9UW53	Q9UW53 aspergillus
24	2243	68.5	659	10 Q41074	Q41074 phaeodactyl
25	2240.5	68.5	664	10 Q39830	Q39830 glycine max
26	2237	68.3	655	10 Q04022	Q04022 arabidopsis
27	2231	68.2	658	10 Q12752	Q12752 phytophthor
28	2229.5	68.1	701	5 Q24928	Q24928 elmeria ten
29	2214.5	67.7	608	10 Q40058	Q40058 hordeum vul
30	2205	67.4	666	10 Q39804	Q39804 glycine max
31	2188.5	66.9	655	3 P87028	P87028 pneumocysti
32	2160.5	66.0	685	3 Q9H601	Q9H601 plicha angu
33	2149.5	65.7	669	10 Q94H00	Q94H00 oryza sativ
34	2139.5	65.4	652	3 Q00053	Q00053 pneumocysti
35	2116.5	64.7	652	5 Q07615	Q07615 plasmodium
36	2108	64.4	651	3 Q9UW41	Q9UW41 rhizopus st
37	2093	63.9	659	5 Q9XUJ2	Q9XUJ2 crassostrea
38	2080	63.6	642	3 Q9UW40	Q9UW40 rhizopus st
39	2079	63.5	647	3 Q59855	Q59855 schizosach
40	2076	63.4	661	5 Q77164	Q77164 entamoeba h
41	2074.5	63.4	648	10 Q40980	Q40980 pisum sativ
42	2069	63.2	646	13 Q73885	Q73885 gallus gall
43	2065.5	63.1	647	3 Q94104	Q94104 pneumocysti
44	2064.5	63.1	645	3 Q94106	Q94106 pneumocysti
45	2057	62.8	652	5 Q26924	Q26924 trypanosoma

ALIGNMENTS

RESULT 1
ID Q9UK02 PRELIMINARY: PRT; 639 AA.
AC Q9UK02:
DT 01-MAY-2000 (TRMBLrel. 13, Created)
DT 01-MAY-2000 (TRMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TRMBLrel. 19, Last annotation update)
BT BIP PROTEIN (FRAGMENT).
GN HSP45.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ARTICULAR CARTILAGE;
RA File M.S.;
RT *BIP: an Autoantigen associated with Rheumatoid Arthritis.*;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
DR EMBL: AF188611; AAF13605.1; -
DR HSP: P19120; 3HSC.
DR InterPro: IPR01023; HSP70.
DR Pfam: PF00012; HSP70.1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
KW ATP-binding. 639
FT NON-TER
SQ SEQUENCE 639 AA: 70931 MM: 3421208FF18PDS5D CRC64:

Query Match 100.0%; Score 3273; DB 4; Length 639;
Best Local Similarity 100.0%; Pred. No. 3.6e-157;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MEDKKEDGTVGIDCTTSCVGFKNQVETIANDQGNRTTPSVAFPEGCEGLID 60
|||||

```

Db 1 MEEDKEDVGTAVGIDLTGTTSCVGFKNRVEIANDGNRIITPSVAFTPEGERLIGD 60
Qy 61 AAKNOLTSNPENTVFDKAKRLIGRTWNPDSVOODIKFLPRKYVEKTKRYIOVDIGGGOTK 120
Db 61 AAKNOLTSNPENTVFDKAKRLIGRTWNPDSVOODIKFLPRKYVEKTKRYIOVDIGGGOTK 120
Qy 121 TFAPEEISAMVLTGKKEETAEAVALGKKTVAHVTPAFYFNDQROATFACCTAGLWVMRI 180
Db 121 TFAPEEISAMVLTGKKEETAEAVALGKKTVAHVTPAFYFNDQROATFACCTAGLWVMRI 180
Qy 181 INPTAAALAYGLDKRREGKNTLVFDLGGTFYDSLITDNGVPEVATNGDNLGGEDF 240
Db 181 INPTAAALAYGLDKRREGKNTLVFDLGGTFYDSLITDNGVPEVATNGDNLGGEDF 240
Qy 241 DQVMEHFILKYKKKTGDKVKNRAVOKLREVEKAKRALLSSQOARIEIESFEYGEDF 300
Db 241 DQVMEHFILKYKKKTGDKVKNRAVOKLREVEKAKRALLSSQOARIEIESFEYGEDF 300
Qy 301 SETLTAKEEELNMDLFRSTMKPVQKVLSDLSKSDIDELVVGSGSTRIPKIQOLVKER 360
Db 301 SETLTAKEEELNMDLFRSTMKPVQKVLSDLSKSDIDELVVGSGSTRIPKIQOLVKER 360
Qy 361 FNCKEPRSGINPDEAVAYGAOVAAGVLSGDDTDGLDYLDPYRLTGLIETVGSVMTKLIP 420
Db 361 FNCKEPRSGINPDEAVAYGAOVAAGVLSGDDTDGLDYLDPYRLTGLIETVGSVMTKLIP 420
Qy 421 RNTVPTKKSQIFSTASDNOPTVTKYEGEERPLTKDNHLLGTFDGLGTPARRGVOTE 480
Db 421 RNTVPTKKSQIFSTASDNOPTVTKYEGEERPLTKDNHLLGTFDGLGTPARRGVOTE 480
Qy 481 VFEIDVNGILVTAEDKGTGNKNTTTNDQNKLTREBELRYMVDNAEKFAEEDKKLYER 540
Db 481 VFEIDVNGILVTAEDKGTGNKNTTTNDQNKLTREBELRYMVDNAEKFAEEDKKLYER 540
Qy 541 IDPRNLESGYAVSLKNQIGDKKGLSKSBDKFTMKAEKREMELESQDADIEDPKA 600
Db 541 IDPRNLESGYAVSLKNQIGDKKGLSKSBDKFTMKAEKREMELESQDADIEDPKA 600
Qy 601 KKELEIYOPILSKLYGSAGRPRTGEEDTAEELNNNNH 639
Db 601 KKELEIYOPILSKLYGSAGRPRTGEEDTAEELNNNNH 639

```

RESULT 2

```

ID 09DC41 PRELIMINARY: PRT: 655 AA.
AC 09DC41;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
RE 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
HEAT SHOCK 70KD PROTEIN 5 (GLUCOSE-REGULATED PROTEIN, 78KD).
GN HSPA5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akazawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamashita T.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleschmann W., Gasteirland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga M., Carolini P., de Bonaldo M.F.,
RA Brownstein M.D., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelinich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

```

```

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -I- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
DR EMBL: AK004576; BAB23387.1; -.
DR HSP: P19120; HSP70.
DR MGI: MGI:95835; Hsp70.
DR InterPro: IPR000866; ER_target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70; 1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
KW ATP-binding.
SQ
SEQUENCE 655 AA; 72422 MW; 69E2C0A2C896DEBC CRC64;

```

Query Match 97.7%; Score 3199; DB 11; Length 655;

Best Local Similarity 99.2%; Pred. No. 2e-153;

Matches 626; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

Qy 2 EEDKEDVGTAVGIDLTGTTSCVGFKNRVEIANDGNRIITPSVAFTPEGERLIGD 61
Db 2 EEDKEDVGTAVGIDLTGTTSCVGFKNRVEIANDGNRIITPSVAFTPEGERLIGD 61
Qy 21 EEDKEDVGTAVGIDLTGTTSCVGFKNRVEIANDGNRIITPSVAFTPEGERLIGD 80
Db 21 EEDKEDVGTAVGIDLTGTTSCVGFKNRVEIANDGNRIITPSVAFTPEGERLIGD 80
Qy 62 AAKNOLTSNPENTVFDKAKRLIGRTWNPDSVOODIKFLPRKYVEKTKRYIOVDIGGGOTK 121
Db 62 AAKNOLTSNPENTVFDKAKRLIGRTWNPDSVOODIKFLPRKYVEKTKRYIOVDIGGGOTK 121
Qy 81 AAKNOLTSNPENTVFDKAKRLIGRTWNPDSVOODIKFLPRKYVEKTKRYIOVDIGGGOTK 140
Db 81 AAKNOLTSNPENTVFDKAKRLIGRTWNPDSVOODIKFLPRKYVEKTKRYIOVDIGGGOTK 140
Qy 122 TFAPEEISAMVLTGKKEETAEAVALGKKTVAHVTPAFYFNDQROATFACCTAGLWVMRI 181
Db 122 TFAPEEISAMVLTGKKEETAEAVALGKKTVAHVTPAFYFNDQROATFACCTAGLWVMRI 181
Qy 141 FAREEISAMVLTGKKEETAEAVALGKKTVAHVTPAFYFNDQROATFACCTAGLWVMRI 200
Db 141 FAREEISAMVLTGKKEETAEAVALGKKTVAHVTPAFYFNDQROATFACCTAGLWVMRI 200
Qy 182 NEPTAAALAYGLDKRREGKNTLVFDLGGTFYDSLITDNGVPEVATNGDNLGGEDF 241
Db 182 NEPTAAALAYGLDKRREGKNTLVFDLGGTFYDSLITDNGVPEVATNGDNLGGEDF 241
Qy 201 NEPTAAALAYGLDKRREGKNTLVFDLGGTFYDSLITDNGVPEVATNGDNLGGEDF 260
Db 201 NEPTAAALAYGLDKRREGKNTLVFDLGGTFYDSLITDNGVPEVATNGDNLGGEDF 260
Qy 242 QVMEHFILKYKKKTGDKVKNRAVOKLREVEKAKRALLSSQOARIEIESFEYGEDF 301
Db 242 QVMEHFILKYKKKTGDKVKNRAVOKLREVEKAKRALLSSQOARIEIESFEYGEDF 301
Qy 261 QVMEHFILKYKKKTGDKVKNRAVOKLREVEKAKRALLSSQOARIEIESFEYGEDF 320
Db 261 QVMEHFILKYKKKTGDKVKNRAVOKLREVEKAKRALLSSQOARIEIESFEYGEDF 320
Qy 302 ETLTAKEEELNMDLFRSTMKPVQKVLSDLSKSDIDELVVGSGSTRIPKIQOLVKER 361
Db 302 ETLTAKEEELNMDLFRSTMKPVQKVLSDLSKSDIDELVVGSGSTRIPKIQOLVKER 361
Qy 321 ETLTAKEEELNMDLFRSTMKPVQKVLSDLSKSDIDELVVGSGSTRIPKIQOLVKER 380
Db 321 ETLTAKEEELNMDLFRSTMKPVQKVLSDLSKSDIDELVVGSGSTRIPKIQOLVKER 380
Qy 362 NGRKPSGGINPDEAVAYGAOVAAGVLSGDDTDGLDYLDPYRLTGLIETVGSVMTKLIP 421
Db 362 NGRKPSGGINPDEAVAYGAOVAAGVLSGDDTDGLDYLDPYRLTGLIETVGSVMTKLIP 421
Qy 381 NGRKPSGGINPDEAVAYGAOVAAGVLSGDDTDGLDYLDPYRLTGLIETVGSVMTKLIP 440
Db 381 NGRKPSGGINPDEAVAYGAOVAAGVLSGDDTDGLDYLDPYRLTGLIETVGSVMTKLIP 440
Qy 422 NTVPTKKSQIFSTASDNOPTVTKYEGEERPLTKDNHLLGTFDGLGTPARRGVOTE 481
Db 422 NTVPTKKSQIFSTASDNOPTVTKYEGEERPLTKDNHLLGTFDGLGTPARRGVOTE 481
Qy 441 NTVPTKKSQIFSTASDNOPTVTKYEGEERPLTKDNHLLGTFDGLGTPARRGVOTE 500
Db 441 NTVPTKKSQIFSTASDNOPTVTKYEGEERPLTKDNHLLGTFDGLGTPARRGVOTE 500
Qy 482 TFEIDVNGILVTAEDKGTGNKNTTTNDQNKLTREBELRYMVDNAEKFAEEDKKLYER 541
Db 482 TFEIDVNGILVTAEDKGTGNKNTTTNDQNKLTREBELRYMVDNAEKFAEEDKKLYER 541
Qy 501 TFEIDVNGILVTAEDKGTGNKNTTTNDQNKLTREBELRYMVDNAEKFAEEDKKLYER 560
Db 501 TFEIDVNGILVTAEDKGTGNKNTTTNDQNKLTREBELRYMVDNAEKFAEEDKKLYER 560
Qy 542 DTRNELESYAVSLKNQIGDKKGLSKSBDKFTMKAEKREMELESQDADIEDPKA 601
Db 542 DTRNELESYAVSLKNQIGDKKGLSKSBDKFTMKAEKREMELESQDADIEDPKA 601
Qy 561 DTRNELESYAVSLKNQIGDKKGLSKSBDKFTMKAEKREMELESQDADIEDPKA 620
Db 561 DTRNELESYAVSLKNQIGDKKGLSKSBDKFTMKAEKREMELESQDADIEDPKA 620
Qy 602 KKELEIYOPILSKLYGSAGRPRTGEEDTAE 632
Db 602 KKELEIYOPILSKLYGSAGRPRTGEEDTAE 632
Qy 621 KKELEIYOPILSKLYGSAGRPRTGEEDTAE 651
Db 621 KKELEIYOPILSKLYGSAGRPRTGEEDTAE 651

```

RESULT 3

```

ID 091688 PRELIMINARY: PRT: 658 AA.
AC 091688;

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:04:54 ; Search time 58.8 Seconds
(without alignments)
1207.077 Million cell updates/sec

Title: US-09-806-955a-1

Perfect score: 3273
Sequence: 1 MEEDKEDVGVYDGLGTT.....AGPPTGEEDTAELHHHHH 639

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

11 number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.032802:*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3273	100.0	639	21	AAV90693
2	3225	98.5	633	21	AAV90694
3	3209	98.0	653	15	AAAR3076
4	3182.5	97.2	653	22	AAE12985
5	3182.5	97.2	653	22	AAAB8253
6	3171	96.9	656	14	AAAR4934
7	3168	96.8	666	15	AAAR3076
8	2717	83.0	656	22	ABBS8666
9	2299	70.2	668	21	AAAG1382
10	2299	70.2	691	21	AAAG1381
11	2258.5	69.0	642	21	AAAG39285

12	2258.5	69.0	658	21	AAAG39284	Arabidopsis thaliana
13	2237	68.3	652	21	AAAG39912	Arabidopsis thaliana
14	2174	66.4	685	22	AAAG70749	S cerevisiae apopt
15	2125.5	64.9	687	22	AAAG70875	C albicans apoptos
16	2083	63.6	646	19	AAAW54364	Heat shock cognate
17	2083	63.6	646	20	AAAY17407	Human heat shock c
18	2083	63.6	646	20	AAAY17408	Mouse heat shock c
19	2083	63.6	646	21	AAAB33649	Mouse heat shock p
20	2083	63.6	646	22	AAAB33651	Human heat shock p
21	2083	63.6	646	22	AAE12987	Human Hsp70 family
22	2083	63.6	646	22	AAAB22535	Human heat shock p
23	2083	63.6	646	22	AAAB22938	GFP-Hsc70 fusion p
24	2074.5	63.4	656	18	AAAM01638	Candida albicans h
25	2064	63.1	646	11	AAAR03927	Rat HSP (ratHsp70)
26	2041	62.4	647	21	AAAR03928	Xenopus laevis HSP
27	2039.5	62.3	624	21	AAAB23252	Human Hsp72 (heat
28	2039.5	62.3	641	19	AAAW54367	Human heat shock 7
29	2039.5	62.3	641	21	AAAB23652	Human heat shock p
30	2039.5	62.3	641	22	AAAB23652	Human Hsp70 family
31	2039.5	62.3	641	22	AAAB23652	Human heat shock p
32	2036.5	62.2	646	21	AAAB82534	Human heat shock p
33	2036.5	62.1	633	14	AAAG33604	Arabidopsis thaliana
34	2032.5	62.1	647	20	AAAR43002	Mouse Slp1p homolo
35	2032.5	62.1	647	22	AAAR43002	T. gondii antigen
36	2032	62.1	647	22	AAAR43009	Toxoplasma gondii
37	2030	62.0	651	22	AAAR03930	Gallus gallus HSP
38	2021	61.7	640	21	AAAB23653	Drosophila melanog
39	2017.5	61.6	640	21	AAAR03929	Human heat shock p
40	2014	61.5	642	21	AAAB23650	Homo sapiens HSP (
41	2013	61.5	640	18	AAAM10065	Rat heat shock pro
42	2013	61.5	640	21	AAAB8408	Human heat shock p
43	2013	61.5	677	11	AAAR09418	Hsp70 antigen from
44	2008	61.4	641	14	AAAR43004	Mouse mouse sperm
45	2004.5	61.2	669	11	AAAR03925	T. cruzi HSP (tc70

ALIGNMENTS

AAV90693	1	AAV90693	standard; Protein: 639 AA.
XX	AC	AAV90693;	
XX	DT	29-AUG-2000 (first entry)	
XX	DE	639 residue human immunoglobulin heavy chain binding protein, BIP(78KD).	
XX	KM	Immunoglobulin heavy chain binding protein, BIP(78KD); chondrocyte;	
XX	KM	autoantigen; rheumatoid arthritis; antiarthritic; antirheumatic; p/8.	
OS	XX	Homo sapiens.	
XX	PN	MO200021995-A1.	
XX	PD	20-APR-2000.	
XX	PF	08-OCT-1999; 99MO-GB03316.	
XX	PR	09-OCT-1998; 98GB-0022115.	
XX	PA	(UNLO) KINGS COLLEGE LONDON.	
XX	PI	Panayl GS, Corrigan VM, Bodman-Smith MD, Fife MS, Lanchbury JS;	
XX	DR	WPI: 2000-317942/27.	
XX	DR	N-PSDB: AAA30792.	
XX	PT	New human immunoglobulin heavy chain binding protein and encoding	
XX	PT	polynucleotide, useful for diagnosis and treatment of rheumatoid	
XX	PT	arthritis	

PS Claim 3; Page 23; 53pp; English.

XX This sequence represents a human immunoglobulin heavy chain binding
CC protein, BIP(78KD), having a 639 amino acid sequence. The invention
CC also encompasses a BIP(78KD) protein of 633 amino acids (Y90694).
CC The cDNA encoding BIP(78KD), also referred to as p78 in the
CC specification, was isolated from human chondrocytes (the specialised
CC cells of articular cartilage) and human chondrosarcoma cell lines. The
CC BIP(78KD) cDNA of this invention contains a number of differences
CC compared with the published sequence (Genbank accession number X87949),
CC and has therefore been deposited with Genbank with the accession
CC number AF188611). These differences comprise 6 single nucleotide
CC substitutions and a codon insertion, and result in three amino acid
CC substitutions and an arginine insertion at position 834-836 of the
CC protein. The BIP(78KD) proteins react with antibodies present in the
CC serum of rheumatoid arthritis patients, and is therefore a putative
CC autoantigen for this autoimmune disease. BIP(78KD) is also able to
CC selectively proliferate synovial T-cells from patients with rheumatoid
CC arthritis. BIP(78KD) or peptides derived from the protein are useful as
CC reagents to indicate the presence of rheumatoid arthritis, and can be
CC used in prognostic or diagnostic tests of body fluids for rheumatoid
CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western
CC blotting. The protein or the cDNA encoding it can also be used to test
CC for rheumatoid arthritis by detecting antibodies to the protein.
CC BIP(78KD), its peptides and polynucleotides are also useful
CC therapeutically.

XX Sequence 639 AA:

Query Match 100.0%; Score 3273; DB 21; Length 639;

Best Local Similarity 100.0%; Pred. No. 1.0e-209; Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MEEEDKEDVGVGIDGIDGTTGSCVGEKNGREITANOGNRIIPSYVAFPEGERILGD 60
DB 1 meekkedvgvlgidgltgtycevgfkngryellandqgnrlpsyaifpgerilgd 60
OY 61 AAKQQLSNENRYFDAKRLIGRTWNPVODIFLCPKRYEKKTKRYIOVDIGGQTK 120
DB 61 aakqqlsnenryfdaakrligrtwnpvovdifiklpkryektkryiovdigggqtk 120
OY 121 TFAPEEISAAVLTAKKETAAYAGKRVTHAVVTPAVFENDARQATKACTAGLANMRI 180
DB 121 tfapeeisaaavltaakketayagkrvthavvtpavfendarqatkactaglanmri 180
OY 121 tfapeeisaaavltaakketayagkrvthavvtpavfendarqatkaglanmri 180
DB 121 tfapeeisaaavltaakketayagkrvthavvtpavfendarqatkaglanmri 180
OY 181 INEPTAAATAYGIDKREGEKNILVFDLGGSTFDVSLITIDKGVFEVATNDOTHLGGE 240
DB 181 ineptaaataygidkregeknilvfdlggstfdvslitidkgvfevatndotlhlgge 240
OY 241 DQRYMEHFIILYKKRTGKDVAKNDRAVQKLREVEKAKRALSSQHARIEIESFEDE 300
DB 241 dqrymehfiiylkkrctgkdvakndravqklrevekakralssqharieiesfede 300
OY 301 SEITLRAKFEELKMDLFRSTMKPVQKVLSDLSKSDIDEIVLVGSGTRIPKIDQLYKE 360
DB 301 seitlrakfeelkmdlfrstmkpvqkvlstdlsksdideivlvvgstripkidlqlyke 360
OY 361 PNGKEPSRGINPDDEAVAYGAOAGVLSGDOQTGLVLDVCPGLTGLGTYGGMATLIP 420
DB 361 pngkepsrginpddeavaygaagvlsdgoqtdglvldvcplgtglgtyggmatlip 420
OY 421 RNTVVPKRSKQIFSTASDNQPYVTIKVYEGEPLTRKDNHLGTFDPLTGTPAPRGVQ 480
DB 421 rntvvpkrskqifstasdnqpyvtikvyegepltrkdnhlgtfdpltgtpaprgvq 480
OY 481 VTFEIDVNGILRTAYAEKDGKGNKNTITINDQNRLTPEIEHMYADAKFAEDKKLKER 540
DB 481 vtfeidvngilrtayaeekdgkgnkntitindqnrltpeiehmnyadakefedkkker 540
OY 541 IDTNELESTAYSLKNOICDEKLGKSSDEKETMEKAVEKEITMLESODADIEDPKA 600
DB 541 idtnelestayslknoicdeklgkssdeketmekavekeitmlesodadiedpka 600

```

OY 601 KKKELEIVOPTITSKLYGSAGAPPTGEEDTAELHHNNHH 639
DB 601 kkeleivoptitsklygsagppptgeedtaelhhnnhh 639

RESULT 2

ID AAY90694 standard; protein; 633 AA.

AC AAY90694;

DT 29-AUG-2000 (first entry)

DE 633 residue human immunoglobulin heavy binding protein, BIP(78KD).

KW Immunoglobulin heavy chain binding protein; BIP(78KD); chondrocyte;
KW autoantigen; rheumatoid arthritis; antirheumatic; p78.

OS Homo sapiens.

PN W0200021995-A1.

PD 20-APR-2000.

PE 08-OCT-1999; 99MO-GB03316.

PR 09-OCT-1998; 98GB-0022115.

PA (UNLO) KINGS COLLEGE LONDON.

PI Parayl GS, Corrigan VM, Bodman-Smith MD, Fife MS, Lancbury JS;

DR WPI; 2000-317942/27.

PT New human immunoglobulin heavy chain binding protein and encoding
PT polynucleotide, useful for diagnosis and treatment of rheumatoid
PT arthritis -

PS Claim 3; Page 44-46; 53pp; English.

XX This sequence represents a human immunoglobulin heavy chain binding
CC protein, BIP(78KD), having a 633 amino acid sequence. The invention
CC also encompasses a BIP(78KD) protein of 639 amino acids (Y90694).
CC The cDNA encoding BIP(78KD), also referred to as p78 in the
CC specification, was isolated from human chondrocytes (the specialised
CC cells of articular cartilage) and human chondrosarcoma cell lines. The
CC BIP(78KD) cDNA of this invention contains a number of differences
CC compared with the published sequence (Genbank accession number X87949),
CC and has therefore been deposited with Genbank with the accession
CC number AF188611). These differences comprise 6 single nucleotide
CC substitutions and a codon insertion, and result in three amino acid
CC substitutions and an arginine insertion at position 834-836 of the
CC protein. The BIP(78KD) proteins react with antibodies present in the
CC serum of Rheumatoid arthritis patients, and is therefore a putative
CC autoantigen for this autoimmune disease. BIP(78KD) is also able to
CC selectively proliferate synovial T-cells from patients with rheumatoid
CC arthritis. BIP(78KD) or peptides derived from the proteins are useful as
CC reagents to indicate the presence of rheumatoid arthritis, and can be
CC used in prognostic or diagnostic tests of body fluids for rheumatoid
CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western
CC blotting. The protein or the cDNA encoding it can also be used to test
CC for rheumatoid arthritis by detecting antibodies to the protein.
CC BIP(78KD), its peptides and polynucleotides are also useful
CC therapeutically.

SO Sequence 633 AA;

Query Match 98.5%; Score 3225; DB 21; Length 633;

Best Local Similarity 100.0%; Pred. No. 2.4e-206; Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:18:33 ; Search time 34.99 Seconds

(without alignments)
1738.341 Million cell updates/sec

Title: US-09-806-955a-2

Perfect score: 3225
Sequence: 1 MEEDKEDVGTVGIDLTGTT.....SKLYSGAPPEPTGEDTAEI 633Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	99.5	654	2 A27414	dnak-type molecule
2	3201	99.3	655	2 A37048	dnak-type molecule
3	3200	99.2	654	1 HHRGB	dnak-type molecule
4	3182.5	98.7	653	2 A29821	dnak-type molecule
5	3168	98.2	652	2 T50242	dnak-type molecule
6	2722.5	84.4	667	2 S24782	dnak-type molecule
7	2719.5	84.3	667	2 D44261	dnak-type molecule
8	2710	84.0	656	2 JN0666	dnak-type molecule
9	2618.5	81.2	661	2 T15513	heat shock 70K pro
10	2598.5	80.4	661	2 A32475	heat shock 70K pro
11	2555	79.2	657	2 T34037	heat shock 70K pro
12	2354.5	73.0	667	2 S21879	dnak-type molecule
13	2346.5	72.8	668	2 S21880	dnak-type molecule
14	2332.5	72.3	668	2 T04080	dnak-type molecule
15	2325	72.1	668	2 T46574	dnak-type molecule
16	2322.5	72.0	663	2 T04078	dnak-type molecule
17	2305.5	71.5	663	2 T03581	dnak-type molecule
18	2295	71.2	668	2 S71171	dnak-type molecule
19	2283	70.8	672	2 T43723	dnak-type molecule
20	2276	70.6	672	2 T43716	dnak-type molecule
21	2261	70.1	642	2 T45298	dnak-type molecule
22	2237	69.4	655	2 H86222	hypothetical prote
23	2234.5	69.3	664	2 T06598	dnak-type molecule
24	2231	69.2	658	2 S38890	dnak-type molecule
25	2227.5	69.1	662	2 T50464	glucose-regulated
26	2216	68.7	663	2 T38155	78 kd glucose regu
27	2214.5	68.7	608	2 T05741	dnak-type molecule
28	2200	68.2	666	2 T06358	dnak-type molecule
29	2173	67.4	682	1 HHRK2	dnak-type molecule

ALIGNMENTS

30	2156	66.9	663	2 S20877	dnak-type molecule
31	2108	65.4	651	2 JC7132	heat shock protein
32	2100	65.1	655	2 A48468	dnak-type molecule
33	2083	64.6	646	2 S07197	dnak-type molecule
34	2083	64.6	646	2 JC4853	dnak-type molecule
35	2083	64.6	646	2 A27027	dnak-type molecule
36	2080	64.5	641	2 PC7036	heat shock protein
37	2079	64.5	647	2 T41121	heat shock protein
38	2079	64.5	650	2 S1456	dnak-type molecule
39	2078	64.4	645	2 S41372	dnak-type molecule
40	2077	64.4	646	2 A45935	dnak-type molecule
41	2076.5	64.4	640	2 S37394	dnak-type molecule
42	2074.5	64.3	648	2 S53498	dnak-type molecule
43	2074.5	64.3	656	2 S51712	dnak-type molecule
44	2074	64.3	646	2 A35922	dnak-type molecule
45	2072	64.2	639	2 S20139	dnak-type molecule

RESULT 1

A27414

dnak-type molecular chaperone GRP78 precursor - Chinese hamster

M:Alternate names: glucose-regulated 78K protein; Ig heavy chain-binding protein

C:Species: Cricetus griseus (Chinese hamster)

C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 20-Aug-1999

C:Accession: A27414

R:Ring, J.; Wooden, S.K.; Kriz, R.; Kelleher, K.; Kaufman, R.J.; Lee, A.S.

Gene 55, 147-152, 1987

A:Title: The nucleotide sequence encoding the hamster 78-kDa glucose-regulated protei

A:Reference number: A27414; MUID:87305586

A:Accession: A27414

A:Molecule type: mRNA

A:Residues: 1-654 <TIN>

A:Cross-references: GB:M17169; NID:g191090; PIDN:AAA51448.1; PID:g304510

C:Genetics:

A:Gene: GRP78

C:Function:

A:Description: involved in protein folding and assembling/dissassembling of protein co

C:Superfamily: heat shock protein 70

C:Keywords: ATP; heterotetramer; immunoglobulin; molecular chaperone

F:1-18/Domain: signal sequence #status predicted <SIC>

F:19-654/Product: Ig heavy chain binding protein #status predicted <MAN>

F:651-654/Region: endoplasmic reticulum retention signal

Query Match	Best Local Similarity	99.5%	Score 3209;	DB 2:	Length 654;
Matches 629;	Conservative	2;	Mismatches 0;	Indels 0;	Gaps 0;
OY	2	MEEDKEDVGTVGIDLTGTTSCVCFKNGREVEIANDQGNRTPSVATPGEERLIGDA	61		
DB	20	EEEDKEDVGTVGIDLTGTTSCVCFKNGREVEIANDQGNRTPSVATPGEERLIGDA	79		
OY	62	ANQUTSPENTVPDAKRLIGRTWNPDSVOODIKFLPFYVEKKTKPYIQVDIGGGQRTK	121		
DB	80	ANQUTSPENTVPDAKRLIGRTWNPDSVOODIKFLPFYVEKKTKPYIQVDIGGGQRTK	139		
OY	122	FAPEISAMVLTGKMETAEVAGKRVHVVTPAVFENDQROATDAGTIGLVNMRIL	181		
DB	140	FAPEISAMVLTGKMETAEVAGKRVHVVTPAVFENDQROATDAGTIGLVNMRIL	199		
OY	182	NEPTAAATAYGLDKREGKNTLVFDLGGTFVSLTINDGVEVYATNGDHLGGEPD	241		
DB	200	NEPTAAATAYGLDKREGKNTLVFDLGGTFVSLTINDGVEVYATNGDHLGGEPD	259		
OY	242	QVMEHFITKLYKKTKGDKVRKNRAVOKLRREYKAKRALSSQHQARIEISFYEGEDFS	301		
DB	260	QVMEHFITKLYKKTKGDKVRKNRAVOKLRREYKAKRALSSQHQARIEISFYEGEDFS	319		
OY	302	ETLTAKREELNMLDFSTAKPVQKVLIEDSLKSDIDELVVGSGTRPKIQOLVKKFF	361		

Db 320 ETLTRAKFELNMDLFRSTMKPVQKVLSDLSKSDIDELVYLGSGSTRIPKIQOLYKEFF 379
 362 NCKEPRSGINPDEAAVAGAAVAGVLSGDDTGDVLVLDVCPRLTGLETIVGGMTKLIPR 421
 Db 380 NCKEPRSGINPDEAAVAGAAVAGVLSGDDTGDVLVLDVCPRLTGLETIVGGMTKLIPR 439
 Qy 422 NTVPPTKKSQIFSTASDNOPTVTIKVYEGEERPLTKDNHLLGTFDLGIIPAPGVQIY 481
 Db 440 NTVPPTKKSQIFSTASDNOPTVTIKVYEGEERPLTKDNHLLGTFDLGIIPAPGVQIY 499
 Qy 482 TFEIDVNGILRTAEDKGTGKNKTTITNDQNRLLTPEIERMWNDAKFAEEDKTKLEKI 541
 Db 500 TFEIDVNGILRTAEDKGTGKNKTTITNDQNRLLTPEIERMWNDAKFAEEDKTKLEKI 559
 Qy 542 DTRNLESTAYSLKNOIGDKELGKLSSEDEKEMEKAVEKEIEMLESHODADIEDFKAK 601
 Db 560 DTRNLESTAYSLKNOIGDKELGKLSSEDEKEMEKAVEKEIEMLESHODADIEDFKAK 619
 Qy 602 KKELEIVQPIITSKLYGSGAPPTGEEDTSE 632
 620 KKELEIVQPIITSKLYGSGAPPTGEEDTSE 650

RESULT 2

A37048

dnak-type molecular chaperone grp78 precursor - mouse

N:Alternate names: B1P: glucose-regulated 78k protein; Ig heavy chain-binding protein

C:Species: Mus musculus (house mouse)

C:Date: 15-Jun-1993 #sequence_revision 15-Jan-1993 #text_change 20-Aug-1999

A:Accession: A37048; B37048; J00094; I49137; A31934

R:Kozakusmtl, Y.; Normington, K.; Press, E.; Slaughter, C.; Sambrook, J.; Gething, M.J.

J. Cell Sci. Suppl. 11, 115-137, 1989

A:Title: Identification of immunoglobulin heavy chain binding protein as glucose-regulat

A:Reference number: A37048; MUID:90130686

A:Accession: A37048

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-655 <K21>

A:Accession: B37048

A:Molecule type: protein

A:Residues: 20-41, 'X', 43-46 <K22>

R:Parfett, C.L.J.; Hofbauer, R.; Brudynaki, K.; Edwards, D.R.; Denhardt, D.T.

Gene 82, 291-303, 1989

A:Title: Differential screening of a cDNA library with cDNA probes amplified in a hetero

A:Reference number: J00094; MUID:90060818

A:Accession: J00094

A:Molecule type: mRNA

A:Residues: 488-590, 'G', 592-595, 'E', 597-655 <PAR>

A:Cross-references: GB:M30779; NID:g193644; PIDN:AAA3742.1; PID:g193645

R:Jman, J.B.; Mote, P.L.; Walford, R.L.; Spindler, S.R.

Gene 158, 225-229, 1995

A:Title: Structure and regulation of the mouse GRP78 (B1P) promoter by glucose and calci

A:Reference number: I49137; MUID:9531621

A:Accession: I49137

A:Status: translation not shown; translated from GB/EMBL/DBBJ

A:Molecule type: DNA

A:Residues: 1-28 <KES>

A:Cross-references: EMBL:U16277; NID:g829364; PIDN:AAA76734.1; PID:g829365

R:Haas, I.G.; Mo, T.

Proc. Natl. Acad. Sci. U.S.A. 85, 2250-2254, 1988

A:Title: cDNA cloning of the immunoglobulin heavy chain binding protein.

A:Reference number: A31934; MUID:88176922

A:Accession: A31934

A:Molecule type: mRNA

A:Residues: 514-589, 'D', 591-595, 'E', 597-655 <HAA>

A:Cross-references: GB:M19351; NID:g192231; PIDN:AAA37315.1; PID:g387113

C:Genetics:

A:Gene: grp78

C:Function:

A:Description: Involved in protein folding and assembling/disassembling of protein comp

C:Superfamily: heat shock protein 70

C:Keywords: ATP; endoplasmic reticulum; glycoprotein; heterotetramer; molecular chaperon

F:20-655/Product: dnak-type molecular chaperone grp78 #status experimental <MAT>
 F:652-655/Region: endoplasmic reticulum retention signal

Query Match 99.3%; Score 3201; DB 2; Length 655;
 Best Local Similarity 99.4%; Pred. No. 1,1e-154;
 Matches 627; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EPDKKEDVGVGIDLGTTISGVGKNGREILLANOGNRIPSPVAFPDEGERLIGA 61
 Db 21 EEDKEDVGVGIDLGTTISGVGKNGREILLANOGNRIPSPVAFPDEGERLIGA 80
 Qy 62 AKNQLTNPENVPFADAKRLIGRTMNDPSVQODIKFLPPKVEKKTPYIOVDIGGQTKT 121
 Db 81 AKNQLTNPENVPFADAKRLIGRTMNDPSVQODIKFLPPKVEKKTPYIOVDIGGQTKT 140
 Qy 122 FAREEISAAVLTAKMETEAVYLGKVTAAVTPAYFNDRATDAGTLAGLNMRT 181
 Db 141 FAREEISAAVLTAKMETEAVYLGKVTAAVTPAYFNDRATDAGTLAGLNMRT 200
 Qy 182 NEPTAAAIAYGLDKREGKNIIVFDLGGTFVSLTTIDNGVEVVAATNDTHLGGEDFD 241
 Db 201 NEPTAAAIAYGLDKREGKNIIVFDLGGTFVSLTTIDNGVEVVAATNDTHLGGEDFD 260
 Qy 242 QRYMEHFILYKKKTKGDKVKNRAVOKLREVEKARALSSQARIEIESFYEGEDRS 301
 Db 261 QRYMEHFILYKKKTKGDKVKNRAVOKLREVEKARALSSQARIEIESFYEGEDRS 320
 Qy 302 ETLTRAKFELNMDLFRSTMKPVQKVLSDLSKSDIDELVYLGSGSTRIPKIQOLYKEFF 361
 Db 321 ETLTRAKFELNMDLFRSTMKPVQKVLSDLSKSDIDELVYLGSGSTRIPKIQOLYKEFF 380
 Qy 362 NCKEPRSGINPDEAAVAGAAVAGVLSGDDTGDVLVLDVCPRLTGLETIVGGMTKLIPR 421
 Db 381 NCKEPRSGINPDEAAVAGAAVAGVLSGDDTGDVLVLDVCPRLTGLETIVGGMTKLIPR 440
 Qy 422 NTVPPTKKSQIFSTASDNOPTVTIKVYEGEERPLTKDNHLLGTFDLGIIPAPGVQIY 481
 Db 441 NTVPPTKKSQIFSTASDNOPTVTIKVYEGEERPLTKDNHLLGTFDLGIIPAPGVQIY 500
 Qy 482 TFEIDVNGILRTAEDKGTGKNKTTITNDQNRLLTPEIERMWNDAKFAEEDKTKLEKI 541
 Db 501 TFEIDVNGILRTAEDKGTGKNKTTITNDQNRLLTPEIERMWNDAKFAEEDKTKLEKI 560
 Qy 542 DTRNLESTAYSLKNOIGDKELGKLSSEDEKEMEKAVEKEIEMLESHODADIEDFKAK 601
 Db 561 DTRNLESTAYSLKNOIGDKELGKLSSEDEKEMEKAVEKEIEMLESHODADIEDFKAK 620
 Qy 602 KKELEIVQPIITSKLYGSGAPPTGEEDTSE 632
 Db 621 KKELEIVQPIITSKLYGSGAPPTGEEDTSE 651

RESULT 3

HHR7GB

dnak-type molecular chaperone precursor - rat

N:Alternate names: B1P: glucose-regulated 78k protein; Ig heavy chain-binding protein

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 22-Jun-1999

A:Accession: A23948; A60134; A26257

R:Kunro, S.; Pelham, H.R.B.

Cell 46, 291-300, 1986

A:Title: An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated p

A:Reference number: A23948; MUID:86245075

A:Accession: A23948

A:Molecule type: mRNA

A:Residues: 1-654 <MUN>

A:Cross-references: GB:M14050; NID:g203150; PIDN:AAA0817.1; PID:g203151

R:Pedersen, R.C.; Browne, A.C.

Science 236, 188-190, 1987

A:Title: Steroidogenesis-activator polypeptide isolated from a rat Leydig cell tumor.

A:Reference number: A60134; MUID:87177981

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:25:17 ; Search time 19.71 Seconds

(without alignments)
1243.505 Million cell updates/sec

Title: US-09-806-955A-2

Perfect score: 3225

Sequence: 1 MEEDKKEGVGVWIGDICT.....SKLYGAGPPPTGEPTAEU 633

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3216	99.7	654	1 GR78_HUMAN	P11021 homo sapien
2	3209	99.5	654	1 GR78_MESAU	P07823 mesocricetu
3	3205	99.4	653	1 GR78_MOUSE	P20029 mus musculu
4	3200	99.2	654	1 GR78_RAT	P06761 rattus norv
5	3168	98.2	652	1 GR78_CHICK	Q09593 gallus gall
6	3069	95.2	658	1 GR78_XENLA	Q01883 xenopus lae
7	2722.5	84.4	657	1 GR78_APLCA	Q16956 aplysia cal
8	2708	84.0	656	1 HS7C_DROME	P26844 drosophila
9	2594.5	80.4	661	1 HS7C_CAREL	P27420 caenorhabdi
10	2354.5	73.0	667	1 BIP4_TOBAC	Q03684 nicotiana t
11	2346.5	72.8	668	1 BIP4_TOBAC	Q03685 nicotiana t
12	2332.5	72.3	663	1 BIP3_MALIZE	O24581 zea mays (m
13	2328.5	72.2	666	1 BIP_LYCES	P49118 lycopersico
14	2324.5	72.0	668	1 BIP_SPIOLO	Q42434 spinacia ol
15	2322.5	72.0	663	1 BIP2_MALIZE	P24067 zea mays (m
16	2298.5	71.3	669	1 BIP1_ARATH	O01K73 arabidopsis
17	2296	71.2	668	1 BIP2_ARATH	Q38043 arabidopsis
18	2227.5	69.1	662	1 GR78_NEUCR	P78695 neurospora
19	2216	68.7	663	1 GR78_SCHPO	P36604 schizosacch
20	2191	67.9	670	1 GR78_YARLI	O09170 yarrowia li
21	2173	67.4	682	1 GR78_YEAST	P16474 saccharomyc
22	2100	65.1	655	1 GR78_PLAFO	Q05866 plasmodium
23	2085.5	64.7	641	1 HS71_YEAST	P10591 saccharomyc
24	2083	64.6	646	1 HS7C_HUMAN	P11142 homo sapien
25	2083	64.5	646	1 HS7C_MOUSE	P08109 mus musculu
26	2079	64.5	650	1 HS7C_BOVIN	P19120 bos taurus
27	2078	64.4	644	1 HS71_PICAN	P53421 picnia angu
28	2076.5	64.4	640	1 HS7C_DICDI	P36415 dictyosteli
29	2074.5	64.3	655	1 HS71_CANAL	P41797 candida alb
30	2072	64.3	646	1 HS7C_CRIGR	P15378 cricetulus
31	2072	64.2	638	1 HS72_YEAST	P15392 saccharomyc
32	2065.5	64.0	679	1 GR78_KLUVA	P22010 kluyveromyc
33	2065	64.0	648	1 HS71_PUCGR	Q01877 puccinia gr

34	2064	64.0	641	1 HS74_YEAST	P22202 saccharomyc
35	2063.5	64.0	649	1 HS73_YEAST	P09435 saccharomyc
36	2060.5	63.9	646	1 HS70_NEUCR	Q01233 neurospora
37	2060	63.9	643	1 HS71_SCHPO	O10265 schizosacch
38	2059	63.8	645	1 HS70_SOYBN	P26413 glycine max
39	2053.5	63.7	651	1 HS7C_PETRY	P09189 petunia hyb
40	2052	63.6	652	1 HS7D_MANSE	Q04639 manduca sex
41	2049	63.5	651	1 HS70_ONCNY	P08108 oncorhynch
42	2046	63.4	641	1 HS72_BOVIN	Q27865 bos taurus
43	2045	63.4	651	1 HS71_ARATH	P22953 arabidopsis
44	2043.5	63.4	639	1 HS72_HUMAN	P54652 homo sapien
45	2039.5	63.2	641	1 HS71_HUMAN	P08107 homo sapien

ALIGNMENTS

RESULT 1
ID GR78_HUMAN STANDARD: PRT: 654 AA.
AC P11021: Q9NPF1.
DT 01-JUL-1989 (Rel. 11, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 78 kDa glucose-regulated protein precursor (GRP 78) (Immunoglobulin heavy chain binding protein) (BiP) (Endoplasmic reticulum lumenal Ca²⁺ binding protein grp78).
GN HSPA5 OR GRP78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=8828347; PubMed=2840249;
RA Ting J., Lee A.S.;
RT Human gene encoding the 78,000-dalton glucose-regulated protein and its pseudogene: structure, conservation, and regulation.*;
RT DNA 7:275-286(1988).
RN (2)
RP SEQUENCE FROM N.A.
RX TISSUE=Cervical carcinoma;
RA Chao C.C.K.;
RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.
RX TISSUE=Plidoblast;
RA Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.;
RT "grp78 is involved in the quality control of the LDI-receptor.*";
RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE FROM N.A.
RA Bernués-Fajardo A., Llewellyn D.H., Campbell A.K., Errington R.R.;
RT "Sequence differences between human grp78/BiP isolated from HeLa cells and previously reported human sequences.*";
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN (5)
RP SEQUENCE OF 1-25 FROM N.A.
RX MEDLINE=93126087; PubMed=1480470;
RA Chao C.C.K., Lin-Chao S.;
RT A direct-repeat sequence of the human BiP gene is required for A23187-mediated inducibility and an inducible nuclear factor binding.*;
RT Nucleic Acids Res. 20:6481-6485(1992).
RN (6)
RP SEQUENCE OF 22-38.
RX TISSUE=Breast carcinoma;
RA MEDLINE=97295304; PubMed=9150946;
RA Rasmussen R.K., Ji H., Eddes J.S., Moritz R.L., Reid G.E.,
RT Simpson R.J., Dorow D.S.;
RT "Two-dimensional electrophoretic analysis of human breast carcinoma proteins: mapping of proteins that bind to the SH3 domain of mixed lineage kinase MLK2.*";

```

RL Electrophoresis 18:568-598(1997).
RN (17)
RP SEQUENCE OF 19-40.
RC TISSUE-Colon Sarcoma:
RX MEDLINE-9729306; PubMed-9150948;
RA J.H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
RT "A two-dimensional gel database of human colon carcinoma proteins.";
RL Electrophoresis 18:605-613(1997).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL: M19645; AAA52614.1;
EMBL: X87949; CAA61201.1;
DR EMBL: AJ271729; CAB71335.1;
DR EMBL: AF216292; AAF42836.1;
DR EMBL: X59669; CAA42595.1;
DR PIR: A29821; A29821.
DR HSSP: P19120; 3HSC.
DR SWISS-2DPAGE: P11021; HUMAN.
DR PMMA-2DPAGE: P11021;
DR PICT-2DPAGE: P11021;
DR Siera-2DPAGE: P11021;
DR MIM: 136120;
DR InterPro: IPR000886; ER_target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70.1.
DR PRINTS: PR00014; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET; 1.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
DR ATP-binding; Endoplasmic reticulum; Signal.
KW SIGNAL
FT CHAIN 1 18
FT SITE 19 654 78 KDA GLUCOSE-REGULATED PROTEIN.
FT CONFLICT 297 297 MISSING (IN REF. 1 AND 2).
FT CONFLICT 418 418 D -> H (IN REF. 1 AND 2).
FT CONFLICT 439 439 R -> S (IN REF. 1 AND 2).
FT CONFLICT 447 447 K -> N (IN REF. 1 AND 2).
FT SEQUENCE 654 AA; 72333 MW; 59B7D8DB5BC32A00 CRC64;
Query Match 99.7%; Score 3216; DB 1; Length 654;
Best Local Similarity 100.0%; Pred. No. 4; 6e-153;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 260 QRVMEHFIRLKYKKTGKDVKRDNRVQKLRREVEKAKRALSQHQARIETESFYEGEDPS 319
OY 302 ETLTRAKFEELNNDLFRSTMKPVOKVLESDSKSDIDEIYVCGSTRIFKIDOLYKEFP 361
DB 320 ETLTRAKFEELNNDLFRSTMKPVOKVLESDSKSDIDEIYVCGSTRIFKIDOLYKEFP 379
OY 362 NGEPKSGINPDCAVAYGAAGVLSGSDPDGVLVDVCPRLTGLIEYGVWTKLIPR 421
DB 380 NGEPKSGINPDCAVAYGAAGVLSGSDPDGVLVDVCPRLTGLIEYGVWTKLIPR 439
OY 422 NYVVPTRKKSQIFSTASDNDPTVTIYVBSERPLTKDNHLLGTFDLTGIPAPRGVQIEV 481
DB 440 NYVVPTRKKSQIFSTASDNDPTVTIYVBSERPLTKDNHLLGTFDLTGIPAPRGVQIEV 499
OY 482 TFEIDVGLIRVTAEDKGGKGNKKTITFDQNRLLPTEIERMVDARKFAEDKKLKERI 541
DB 500 TFEIDVGLIRVTAEDKGGKGNKKTITFDQNRLLPTEIERMVDARKFAEDKKLKERI 559
OY 542 DTRNELFSYVSLKNOIGKELGKLSSEDEKTEKAVEEIKTMLESKODADIEDPKAK 601
DB 560 DTRNELFSYVSLKNOIGKELGKLSSEDEKTEKAVEEIKTMLESKODADIEDPKAK 619
OY 602 KKELEETVQPIISKLYGSAGPPTGEDTAE 632
DB 620 KKELEETVQPIISKLYGSAGPPTGEDTAE 650
RESULT 2
GR78_MESAU
ID GR78_MESAU STANDARD; PRT; 654 AA.
AC P07823;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 78 kDa glucose-regulated protein precursor (GRP 78) (Immunoglobulin
DE heavy chain binding protein) (BIP).
GN HSPA5 OR GRP78.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_Taxid=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87305586; PubMed-3623104;
RA Ting J., Wooden S.K., Kriz R., Kelleher K., Kaufman R.J., Lee A.S.;
RT "The nucleotide sequence encoding the hamster 78-kDa
RT glucose-regulated protein (GRP78) and its conservation between
RT hamster and rat.";
RL Gene 55:147-152(1987).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL: M17169; AAA51448.1;
DR PIR: A27414; A27414.
DR HSSP: P19120; 3HSC.
DR InterPro: IPR000886; ER_target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70.1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET; 1.
DR PROSITE: PS00297; HSP70_1; 1.

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:23:16 ; Search time 50.75 Seconds

(without alignments)
2157.750 Million cell updates/sec

Title: US-09-806-955a-2

Perfect score: 3225

Sequence: 1 MEEDKEDVGVVGDLTGT.....SKLXGSAQPPPTGEEDPAEL 633

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

● 1 number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.protozoa:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriophage:*
17: sp.archaeoprotein:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3225	100.0	639	4	09UK02	09UK02 homo sapien
2	3199	99.2	635	11	09DC41	09DC41 mus musculu
3	2994	92.8	658	13	091688	091688 xenopus lae
4	2725	84.5	658	5	076180	076180 bombyx mori
5	2717	84.2	656	5	09VYU3	09VYU3 drosophila
6	2607.5	80.9	656	5	062568	062568 suberites d
7	2555	79.2	657	5	002085	002085 caenorhabdi
8	2534	78.6	651	5	024798	024798 echinococcu
9	2525.5	78.3	649	5	024895	024895 echinococcu
10	2460	76.3	648	5	045038	045038 schistosoma
11	2344.5	72.7	665	10	09M4E8	09M4E8 cucumis sat
12	2325.5	72.1	668	10	022639	022639 glycine max
13	2325	72.1	668	10	040924	040924 pseudotsuga
14	2310.5	71.6	663	10	024182	024182 oryza sativ
15	2305.5	71.5	663	10	09FSY7	09FSY7 corylus ave
16	2291	71.0	668	10	09FSY7	09FSY7 corylus ave

17	2288.5	71.0	665	10	0941K4	0941K4 scherffelia
18	2283	70.8	672	3	014453	014453 aspergillus
19	2276	70.6	672	3	013280	013280 aspergillus
20	2273	70.5	668	5	09U540	09U540 toxoplasma
21	2263	70.2	669	3	09C1C1	09C1C1 cronartium
22	2261	70.1	642	5	061001	061001 toxoplasma
23	2244.5	68.6	672	3	09UW23	09UW23 aspergillus
24	2243	68.6	659	10	041074	041074 phaeodactyl
25	2237	68.4	655	10	004022	004022 arabidopsis
26	2234.5	68.3	664	10	039830	039830 glycine max
27	2231	68.2	658	10	012752	012752 phytophthor
28	2226.5	68.0	701	5	024928	024928 elmeria ten
29	2214.5	68.7	608	10	040058	040058 hordeum vul
30	2200	68.2	666	10	039804	039804 glycine max
31	2188.5	67.9	655	3	P87028	P87028 pneumocysti
32	2156.5	66.9	685	3	09H6U1	09H6U1 plicha angu
33	2149.5	66.7	689	10	094H00	094H00 oryza sativ
34	2139.5	66.3	652	3	000053	000053 pneumocysti
35	2116.5	65.6	652	5	007615	007615 plasmodium
36	2108	65.4	651	3	09UW41	09UW41 rhizopus st
37	2093	64.9	659	5	09X2J2	09X2J2 crassostrea
38	2089	64.5	642	3	09UW00	09UW00 rhizopus st
39	2079	64.5	647	3	059855	059855 schizosacch
40	2074.5	64.3	648	10	040980	040980 pistum saliv
41	2074	64.3	661	5	071164	071164 entamoeba h
42	2069	64.2	646	13	073885	073885 gallus gall
43	2063.5	64.0	647	3	094104	094104 pneumocysti
44	2062.5	64.0	645	3	094106	094106 pneumocysti
45	2057	63.8	652	5	026924	026924 trypanosoma

ALIGNMENTS

RESULT 1
ID 09UK02 PRELIMINARY; PRT; 639 AA.
AC 09UK02;
DT 01-MAY-2000 (TREMBLrel. 13. Created)
DT 01-MAY-2000 (TREMBLrel. 13. Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 15. Last annotation update)
DE BIP PROTEIN (FRAGMENT).
GN HSPA5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ARTICULAR CARTILAGE;
RA File M.S.;
RT "BIP: an Autoantigen associated with Rheumatoid Arthritis.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
DR EMBL: AF186611; AAF13605.1; -;
DR HSSP: P19120; 3HSC.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70.1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
KW ATP-binding.
FT NON_TER
SQ SEQUENCE 639 AA: 70931 MW: 3421208FF128FD5D CRC64;

Query Match 100.0%; Score 3225; DB 4; Length 639;
Best Local Similarity 100.0%; Pred. No. 9e-155;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MEEDKEDVGVVGDLTGTTCVGVFKRNGRVELIANDOSNRITPSVAFTGEERLLGD 60
|||||

Db 1 MEDKEDVGTAVGIDLTGTTSCVGFKNRGVEIANDQGNRLTPSVATPGEGRILGD 60
 QY 61 AAKNOLTSNPENTVDAKRLIGRTWNPSPVODIKFLPEKVEKTKPYIOVDIGGQTK 120
 Db 61 AAKNOLTSNPENTVDAKRLIGRTWNPSPVODIKFLPEKVEKTKPYIOVDIGGQTK 120
 QY 121 TFAPEISAMVLTAKMTEAEVLYGKKTAVAVTPAVFENDQROATDAGTIGLVNMR 180
 Db 121 TFAPEISAMVLTAKMTEAEVLYGKKTAVAVTPAVFENDQROATDAGTIGLVNMR 180
 QY 181 INPTAAALATAGLDKREGKNTLVFDLGGTFDVSLLTIDNGVEVATNGDHLGCEP 240
 Db 181 INPTAAALATAGLDKREGKNTLVFDLGGTFDVSLLTIDNGVEVATNGDHLGCEP 240
 QY 241 DQVMEHFILKLYKKTKGKDVKNRAVOKLRREVEKAKRALSQHQARIESFYEGBE 300
 Db 241 DQVMEHFILKLYKKTKGKDVKNRAVOKLRREVEKAKRALSQHQARIESFYEGBE 300
 QY 301 SETLTAKFEELNMDLFRSTMKPVQKLESDLSKSDIDEVLVGGSTRIPKIQOLVKE 360
 Db 301 SETLTAKFEELNMDLFRSTMKPVQKLESDLSKSDIDEVLVGGSTRIPKIQOLVKE 360
 QY 361 FNCKEPRSGINPDEAVAYGAAGVLSGDDTGDVLVDVCPPLTGLIEVGVMTKLIP 420
 Db 361 FNCKEPRSGINPDEAVAYGAAGVLSGDDTGDVLVDVCPPLTGLIEVGVMTKLIP 420
 QY 421 RNTVVPKKSQIFSTASDNQPTVTIKYVDEPRPLKONHLTGTDLIGIPARQVPLE 480
 Db 421 RNTVVPKKSQIFSTASDNQPTVTIKYVDEPRPLKONHLTGTDLIGIPARQVPLE 480
 QY 481 VFEIDVNGILRTVAEDKGTGNKNTITNDONRLTPPEIERVANDAEKFAEDKLLER 540
 Db 481 VFEIDVNGILRTVAEDKGTGNKNTITNDONRLTPPEIERVANDAEKFAEDKLLER 540
 QY 541 IDPRNELESAYSLKNOIGDKKLGKLSSECKTEMEKAVEKTEMLESHODADIEDPK 600
 Db 541 IDPRNELESAYSLKNOIGDKKLGKLSSECKTEMEKAVEKTEMLESHODADIEDPK 600
 QY 601 KKELEBIQPIISIKLYGSAGPPPTGEEDTAEL 633
 Db 601 KKELEBIQPIISIKLYGSAGPPPTGEEDTAEL 633

RESULT 2
 Q9DC41 PRELIMINARY: PRT: 655 AA.
 AC Q9DC41
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 L 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 L 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE HEAT SHOCK 70KD PROTEIN 5 (GLUCOSE-REGULATED PROTEIN, 78KD).
 GN HSPAS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; TISSUE=LUNG;
 RX MEDLINE=21085660; PubMed=11217851.
 RA Kawai J., Sliogawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Saito T., Okazaki Y., Gojouri T., Bono H., Kasukawa T., Saito R.,
 RA Kodota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glisic C., King B., Kochiwa H.,
 RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Stambli F., Suzuki R., Tomita M., Wagner L., Washino T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brønsgtein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guelinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang R.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshwan-Boris A., Yoshida K., Hasegawa Y., Kawai J.H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
 DR EMBL: AK004578; BAB23387.1; -.
 DR HSSP: P19120; 3HSC.
 DR MED: MGI:39835; Hspas.
 DR InterPro: IPR000886; ER_target.
 DR InterPro: IPR01023; HSP70.
 DR Pfam: PF00012; HSP70; 1.
 DR PRINTS: PR00301; HEATSHOCK70.
 DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
 KW ATP-binding
 SO SEQUENCE 655 AA; 72422 MW; 69E2C0A2C896DEBC CRC64;

Query Match 99.2%; Score 3199; DB 11; Length 655;
 Best Local Similarity 99.2%; Pred. No. 1.9e-153;
 Matches 626; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 EEDKEDVGTAVGIDLTGTTSCVGFKNRGVEIANDQGNRLTPSVATPGEGRILGD 61
 Db 21 EEDKEDVGTAVGIDLTGTTSCVGFKNRGVEIANDQGNRLTPSVATPGEGRILGD 80
 QY 62 AAKNOLTSNPENTVDAKRLIGRTWNPSPVODIKFLPEKVEKTKPYIOVDIGGQTK 121
 Db 81 AAKNOLTSNPENTVDAKRLIGRTWNPSPVODIKFLPEKVEKTKPYIOVDIGGQTK 140
 QY 122 TFAPEISAMVLTAKMTEAEVLYGKKTAVAVTPAVFENDQROATDAGTIGLVNMR 181
 Db 141 TFAPEISAMVLTAKMTEAEVLYGKKTAVAVTPAVFENDQROATDAGTIGLVNMR 200
 QY 182 NEPTAAALATAGLDKREGKNTLVFDLGGTFDVSLLTIDNGVEVATNGDHLGCEP 241
 Db 201 NEPTAAALATAGLDKREGKNTLVFDLGGTFDVSLLTIDNGVEVATNGDHLGCEP 260
 QY 242 QVMEHFILKLYKKTKGKDVKNRAVOKLRREVEKAKRALSQHQARIESFYEGBE 301
 Db 261 QVMEHFILKLYKKTKGKDVKNRAVOKLRREVEKAKRALSQHQARIESFYEGBE 320
 QY 302 ETLTAKFEELNMDLFRSTMKPVQKLESDLSKSDIDEVLVGGSTRIPKIQOLVKE 361
 Db 321 ETLTAKFEELNMDLFRSTMKPVQKLESDLSKSDIDEVLVGGSTRIPKIQOLVKE 380
 QY 362 NGKEPRSGINPDEAVAYGAAGVLSGDDTGDVLVDVCPPLTGLIEVGVMTKLIP 421
 Db 381 NGKEPRSGINPDEAVAYGAAGVLSGDDTGDVLVDVCPPLTGLIEVGVMTKLIP 440
 QY 422 NTVVPKKSQIFSTASDNQPTVTIKYVDEPRPLKONHLTGTDLIGIPARQVPLE 481
 Db 441 NTVVPKKSQIFSTASDNQPTVTIKYVDEPRPLKONHLTGTDLIGIPARQVPLE 500
 QY 482 TFEIDVNGILRTVAEDKGTGNKNTITNDONRLTPPEIERVANDAEKFAEDKLLER 541
 Db 501 TFEIDVNGILRTVAEDKGTGNKNTITNDONRLTPPEIERVANDAEKFAEDKLLER 560
 QY 542 DTRNELESAYSLKNOIGDKKLGKLSSECKTEMEKAVEKTEMLESHODADIEDPK 601
 Db 561 DTRNELESAYSLKNOIGDKKLGKLSSECKTEMEKAVEKTEMLESHODADIEDPK 620
 QY 602 KKELEBIQPIISIKLYGSAGPPPTGEEDTAEL 632
 Db 621 KKELEBIQPIISIKLYGSAGPPPTGEEDTAEL 651

RESULT 3
 Q91688 PRELIMINARY: PRT: 658 AA.
 AC Q91688;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:17:07 ; Search time 58.8 seconds
(without alignments)
1195.743 Million cell updates/sec

Title: US-09-806-955a-2

Sequence: 3225
1 MEEDKEDVGVVIGIDLGTT.....SKLYGSAQPPPTGEBDTAL 633

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

1 number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3225	100.0	633	21	AAV90694
2	3225	100.0	639	21	AAV90693
3	3209	99.5	654	15	AAE51075
4	3182.5	98.7	653	22	AAE12965
5	3182.5	98.7	653	22	AAE82533
6	3171	98.3	656	14	AAE4934
7	3168	98.2	656	15	AAE33076
8	2717	84.2	656	22	AAE58666
9	2296	71.2	668	21	AAE41382
10	2296	71.2	691	21	AAE41381
11	2258.5	70.0	642	21	AAE39285

12	2258.5	70.0	658	21	AAE39284	Arabidopsis thalia
13	2257	69.4	655	21	AAE30912	Arabidopsis thalia
14	2173	67.4	682	22	AAE70749	S cerevisiae appt
15	2122.5	65.8	687	22	AAE70875	C albicans apopt
16	2083	64.6	646	19	AAE54364	Heat shock cognate
17	2083	64.6	646	20	AAE17407	Human heat shock
18	2083	64.6	646	20	AAE17408	Mouse heat shock
19	2083	64.6	646	21	AAE23649	Mouse heat shock
20	2083	64.6	646	21	AAE23651	Human heat shock
21	2083	64.6	646	22	AAE12967	Human Hsp70 family
22	2083	64.6	646	22	AAE82535	Human heat shock
23	2083	64.6	890	22	AAE22938	GFP-HSC70 fusion p
24	2074.5	64.3	656	18	AAE01638	Candida albicans h
25	2064	64.0	645	11	AAE03920	Rat HSP (ratHsp70)
26	2041	63.3	647	11	AAE03928	Xenopus laevis HSP
27	2039.5	63.2	624	21	AAE23252	Human Hsp72 (heat
28	2039.5	63.2	641	19	AAE24349	Human heat shock
29	2039.5	63.2	641	21	AAE23652	Human heat shock
30	2039.5	63.2	641	22	AAE12966	Human Hsp70 family
31	2039.5	63.2	641	22	AAE82534	Human heat shock
32	2036.5	63.1	646	21	AAE53604	Arabidopsis thalia
33	2033.5	63.1	633	14	AAE43002	Mouse SLPI homolo
34	2032.5	63.0	647	20	AAE13880	T. gondii antigen
35	2032.5	63.0	647	22	AAE49059	Toxoplasma gondii
36	2032	63.0	634	11	AAE03930	Gallus gallus HSP
37	2030	62.9	651	22	AAE60514	Drosophila melanog
38	2021	62.7	640	21	AAE23653	Human heat shock
39	2017.5	62.6	640	11	AAE03929	Homo sapiens HSP (
40	2014	62.4	642	21	AAE23650	Rat heat shock pro
41	2013	62.4	640	18	AAE10065	Human heat shock p
42	2013	62.4	677	11	AAE88408	Human heat shock
43	2013	62.4	677	11	AAE09418	Hsp70 antigen from
44	2008	62.3	641	14	AAE43004	Mature mouse sperm
45	2004.5	62.2	669	11	AAE03925	T. cruzi HSP (tc70

ALIGNMENTS

RESULT	1
ID	AAV90694 standard; Protein: 633 AA.
XX	AAV90694;
XX	
DT	29-AUG-2000 (first entry)
XX	
DE	633 residue human immunoglobulin heavy binding protein, BIP(78KD).
XX	
KW	Immunoglobulin heavy chain binding protein; BIP(78KD); chondrocyte;
KW	autoantigen; Rheumatoid arthritis; antiarthritic; antirheumatic; p18.
XX	
OS	Homo sapiens.
XX	
PN	W0200021995-A1.
XX	
PD	20-APR-2000.
XX	
PF	08-OCT-1999; 99WO-GB03316.
XX	
PR	09-OCT-1998; 98GB-0022115.
XX	
PA	(UNLO) KINGS COLLEGE LONDON.
PI	Panayi GS, Corrigall VM, Bodman-Smith MD, Fife MS, Lanchbury JS;
XX	
DR	WPI: 2000-317942/27.
XX	
PT	New human immunoglobulin heavy chain binding protein and encoding
PT	polynucleotide, useful for diagnosis and treatment of Rheumatoid
XX	arthritis -
PS	Claim 3: Page 44-46; 53pp; English.

XX This sequence represents a human immunoglobulin heavy chain binding
CC protein, B1P(78Kd), having a 633 amino acid sequence. The invention
CC also encompasses a B1P(78Kd) protein of 639 amino acids (7906953).
CC The cDNA encoding B1P(78Kd), also referred to as p78 in the
CC specification, was isolated from human chondrocytes (the specialised
CC cells of articular cartilage) and human chondrosarcoma cell lines. The
CC B1P(78Kd) cDNA of this invention contains a number of differences
CC compared with the published sequence (Genbank accession number X87949)
CC and has therefore been deposited with Genbank with the accession
CC number AF186111. These differences comprise 5 single nucleotide
CC substitutions and a codon insertion, and result in three amino acid
CC substitutions and an arginine insertion at position 834-836 of the
CC protein. The B1P(78Kd) proteins react with antibodies present in the
CC serum of rheumatoid arthritis patients, and is therefore a putative
CC autoantigen for this autoimmune disease. B1P(78Kd) is also able to
CC selectively proliferate synovial T-cells from patients with rheumatoid
CC arthritis. B1P(78Kd) or peptides derived from the protein are useful as
CC reagents to indicate the presence of rheumatoid arthritis, and can be
CC used in prognostic or diagnostic tests of body fluids for rheumatoid
CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western
CC blotting. The protein or the cDNA encoding it can also be used to test
CC for rheumatoid arthritis by detecting antibodies to the protein.
CC B1P(78Kd), its peptides and polynucleotides are also useful
CC therapeutically.

SQ Sequence 633 AA;

Query Match	100.0%	Score 3225;	DB 21;	Length 633;
Best Local Similarity	100.0%	Pred. No. 2.8e-208;		
Matches 633; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

Qy	1	MEIDKKEDVGTAVGIDLTCTTSCYGVKRNORVELIANDDGNITPFSYAFFPEERLILIGD	60
Dd	1	meedkkedvgtavgidltcttscygvkrnorvelilanddgnitpfsyafpeeerliligd	60
Qy	61	AAKQUTSIPENTVEDAKRLIGRTWMDPSVQODIKFLPKVVEKTKRYEIVDIGGQTK	120
Dd	61	aaknqtisnpentvedakrligrtwmdpsvqodikflpkvvektkryeivdiggsqtk	120
Qy	121	TKAPEEISAMVLTJMKKEETAEAYLGCKRTTHAVVVPVAFYFNDAOROKTPOAGTAGJANVRI	180
Dd	121	tkapeeisamvltjmkkeetaeaylgckrtthavvvpvafyfnadaoroktpoagtagjanvri	180
Qy	181	INEPTAAATAGLDRKEBEKNILVFDLGGTFDVSLLTTDNGVFEVYATNGSDTHLSEDF	240
Dd	181	ineptaaatagldrkebeknilvfdlggtfdvslslltdngvfevayatngsdthlsedf	240
Qy	241	DQWMEHFIKLRKKTGTGDKRDKRRAQOKRRREKAKALSSQOARIEIESFEEGEDE	300
Dd	241	dqwmehfiklrkktgtgdkrdkrraokrrreekakalssqoarieiesfeegede	300
Qy	301	SELTLTAKEEELMNDLFSTYMKPQVQKLEBSDLAKSDIDBIYLVGSGSTRIPKIQULYKEF	360
Dd	301	seltltakeeelmndlfstymkpqvqklesbdlakstdidbiylvsgstripkiiqulykef	360
Qy	361	FNCKEFSRQIINPDEAVAAQAVQAGVLSGQDITGDVLVLDVCPRLIGETPGVWTLIP	420
Dd	361	fnckefsrqiiinpdeavaaqaavqagvlsqgditgdvlvldvcpirligetpgvwtlrip	420
Qy	421	RNVYVPRKKSQIFSTASDNOPTVITIKYBEERPLUTDNHLLGTFDULGTGIFAPRAGVQIE	480
Dd	421	rnyvprkksqifstasdnoptvitikeybeerplutdnhllgtdfduigtgifapragvqie	480
Qy	481	VFEFIDVNOILRNTAEDGAGTGNKNKIIYTDQNRLLRPEETERNVNDAKFAFAPDKLKER	540
Dd	481	vfefidvnoilrntaedgagtnknkiiytdqnrlrpeeternvndakfafapdklker	540
Qy	541	IDPRNELSEYATSLKNOIGDKKELGKRLSEDEETHEKAAVEKEIEMLESNODADIEDPKA	600
Dd	541	idprnelseyatyslknogdkkelgkrlsedeeethekaaevekeiemlesnodadiedpka	600

QY	601	KKKELEIVOPISKLYGSAGPPPTGEEDTAEL	6333
Db	601	KKKELEELVQPLISKLYGSAGPPPTGEEDTAEL	6333

ID AAY90693 standard; Protein; 639 AA.

AC AAY90693;

DT 29-AUG-2000 (first entry)

DE 639 residue human immunoglobulin heavy binding protein, B1P(78KD).

KW Immunoglobulin heavy chain binding protein; BiP (78KD); chondrocyte;
KW autoantigen; rheumatoid arthritis; antiarthritic; antirheumatic; p78

OS Homo sapiens.

PN W0200021995-A1.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-GB03316.

PR 09-OCT-1998; 98GB-0022115.

PA (UNLO) KINGS COLLEGE LONDON.

PI Panayl GS, Corriganl VM, Bodman-Smith MD, Flfe MS, Lanchbury JS;

DR WPI; 2000-317942/27.

XX
XX

PT polynucleotide, useful for diagnosis and treatment of rheumatoid

[illegible]

CC This sequence represents a human immunoglobulin heavy chain binding
CC protein, B1P(78KD), having a 639 amino acid sequence. The invention
CC also encompasses a B1P(78KD) protein of 633 amino acids (Y90594).
CC The cDNA encoding B1P(78KD), also referred to as p78 in the
CC specification, was isolated from human chondrocytes (the specialised
CC cells of articular cartilage) and human chondrosarcoma cell lines. The
CC B1P(78KD) cDNA of this invention contains a number of differences
CC compared with the published sequence (Genbank accession number X87949),
CC and has therefore been deposited with Genbank with the accession
CC number AF188611). These differences comprise 6 single nucleotide
CC substitutions and a codon insertion, and result in three amino acid
CC substitutions and an arginine insertion at position 834-836 of the
CC protein. The B1P(78KD) proteins react with antibodies present in the
CC serum of rheumatoid arthritis patients, and is therefore a putative
CC autoantigen for this autoimmune disease. B1P(78KD) is also able to
CC selectively proliferate synovial T-cells from patients with rheumatoid
CC arthritis. B1P(78KD) or peptides derived from the protein are useful as
CC reagents to indicate the presence of rheumatoid arthritis, and can be
CC used in prognostic or diagnostic tests of body fluids for rheumatoid
CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western
CC blotting. The protein or the cDNA encoding it can also be used to test
CC for rheumatoid arthritis by detecting antibodies to the protein.
CC B1P(78KD), its peptides and polynucleotides are also useful
CC therapeutically.

sq Sequence 639 AA;

Query Match	100.08;	Score 3225;	DB 21;	Length 639;
Best Local Similarity	100.08;	Pred. No. 2.8e+208;		
Matches 633; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:17:45 ; Search time 26.02 seconds
(without alignments)
594.212 Million cell updates/sec

Title: US-09-806-955a-2

Perfect score: 3225

Sequence: 1 MEDKKEDVGTWVGIDLT.....SKLYGSAGPPPTGEEDTAL 633

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2.6/prodata/2/1aa/5A.COMB.pep: *
2: /cgn2.6/prodata/2/1aa/5B.COMB.pep: *
3: /cgn2.6/prodata/2/1aa/6A.COMB.pep: *
4: /cgn2.6/prodata/2/1aa/6B.COMB.pep: *
5: /cgn2.6/prodata/2/1aa/6C.COMB.pep: *
6: /cgn2.6/prodata/2/1aa/6D.COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	99.5	654	1	US-08-441-139-11
2	3168	98.2	666	1	US-08-441-139-16
3	2173	67.4	682	1	US-08-441-139-2
4	2156	66.9	663	1	US-08-441-139-7
5	2077	64.4	646	1	US-08-441-139-14
6	2063.5	64.0	679	1	US-08-441-139-5
7	1938	60.1	643	4	US-08-441-139-3
8	1547.5	48.0	679	1	US-08-214-583-2
9	1474	45.7	641	1	US-08-441-139-4
10	1461.5	45.3	649	4	US-09-066-047-5
11	1444	44.8	607	2	US-08-472-534-5
12	1303.5	40.4	600	6	5240706-1
13	1140.5	35.4	339	2	US-08-928-692-52
14	941	29.2	187	6	5196523-13
15	824	25.6	168	1	US-08-441-139-10
16	818.5	25.4	165	1	US-08-441-139-10
17	750.5	23.3	471	1	US-08-203-905B-7
18	726.5	22.5	472	1	US-08-203-905B-14
19	701	21.7	307	4	US-08-858-207A-481
20	642.5	19.9	139	6	5196523-10
21	607.5	18.8	999	2	US-08-770-301A-3
22	607.5	18.8	999	2	US-09-175-581-3
23	598	18.5	999	3	US-08-770-301A-1
24	528.5	16.4	560	3	US-09-175-581-1
25	528.5	16.4	560	3	US-08-928-692-53
26	437.5	13.6	374	2	US-08-928-692-51
27	402	12.5	77	6	5196523-7

28	395	12.2	79	6	5196523-11	Patent No. 5196523
29	367.5	11.4	599	4	US-09-080-983-9	Sequence 9, App11
30	358	11.1	549	2	US-08-770-544-6	Sequence 6, App11
31	345.5	10.7	80	1	US-08-464-164-4	Sequence 4, App11
32	345.5	10.7	80	1	US-08-338-057-4	Sequence 4, App11
33	345.5	10.7	80	2	US-08-668-416-4	Sequence 4, App11
34	323	10.0	554	5	PCT-US94-06430-7	Sequence 7, App11
35	320	9.9	554	4	US-08-591-468-7	Sequence 7, App11
36	267.5	8.3	714	2	US-08-472-534-3	Sequence 7, App11
37	264	8.2	36	6	5196523-12	Sequence 3, App11
38	230	7.1	46	6	5196523-8	Patent No. 5196523
39	211	6.5	42	6	5196523-15	Patent No. 5196523
40	192	6.0	38	6	5196523-9	Patent No. 5196523
41	168.5	5.2	1786	4	US-08-973-462-8	Sequence 8, App11
42	166	5.1	125	4	US-08-858-207A-407	Sequence 407, App
43	147	4.6	693	4	US-08-235-836C-68	Sequence 68, App11
44	143.5	4.4	1805	1	US-07-853-913-2	Sequence 2, App11
45	142.5	4.4	2285	4	US-09-308-373-2	Sequence 2, App11

ALIGNMENTS

RESULT 1
US-08-441-139-11
Sequence 11, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: Wiltrop, Dr. Karl D.
APPLICANT: Robinson, Anne S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-11
Query Match 99.5%; Score 3209; DB 1; Length 654;
Best Local Similarity 99.7%; Pred. No. 3.5e-259;
Matches 629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 EEDKKEDVGTAVGIDLTGTYSCVGFKNKGRVEIITANDGNRITPSYAFTPEGERLIGDA 61
DB 20 EEDKKEDVGTAVGIDLTGTYSCVGFKNKGRVEIITANDGNRITPSYAFTPEGERLIGDA 79
QY 62 AKNQLTSPNENTVFDAKRLIGRTWMDPSVQODIKFLPKRYVEKTKPYIOVDIGGCGT 121
DB 80 AKNQLTSPNENTVFDAKRLIGRTWMDPSVQODIKFLPKRYVEKTKPYIOVDIGGCGT 139
QY 122 FAPBEISAMVLTAKMTAEAYLGKVTYHAYVTPAYFNDAROKRTKAGTACIAGLNMWRII 181
DB 140 FAPBEISAMVLTAKMTAEAYLGKVTYHAYVTPAYFNDAROKRTKAGTACIAGLNMWRII 199
QY 182 NEPTAAIAVGLDKREGKENILVFDLGGCTFDVSLITDNGVFEVATNGDTHLGEDPD 241
DB 200 NEPTAAIAVGLDKREGKENILVFDLGGCTFDVSLITDNGVFEVATNGDTHLGEDPD 259
QY 242 ORVMEHFIRLYKKTKGDKVKNRAVOKLRREVEKAKRALSQHOARIEIESFEYEGEDFS 301
DB 260 ORVMEHFIRLYKKTKGDKVKNRAVOKLRREVEKAKRALSQHOARIEIESFEYEGEDFS 319
QY 302 ETLTRAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLVGSTRIRKIOOLYKEFP 361
DB 320 ETLTRAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLVGSTRIRKIOOLYKEFP 379
QY 362 NGKEPSKGINPDEAVAYGAAGVLSGDODTGDVLVDVCPPLTGLGIEYGVWTKLIIPR 421
DB 380 NGKEPSKGINPDEAVAYGAAGVLSGDODTGDVLVDVCPPLTGLGIEYGVWTKLIIPR 439
QY 422 NTVPPTKKSQIFSTASDNPVTIIVYEGSERPLTKDNHLLCTFDLTIIPAPRGVPOIEV 481
DB 440 NTVPPTKKSQIFSTASDNPVTIIVYEGSERPLTKDNHLLCTFDLTIIPAPRGVPOIEV 499
QY 482 TFEIDVNGILRYTAEDEKGTGNKNKITITNDONRLTPEIEEMVNDAKFAEEDKKLKERI 541
DB 500 TFEIDVNGILRYTAEDEKGTGNKNKITITNDONRLTPEIEEMVNDAKFAEEDKKLKERI 559
QY 542 DTRNELESYAVSLKNOIGKEKLGKLSSEDEKTEMEKAVEKIEWLESHODADIEDPKAK 601
DB 560 DTRNELESYAVSLKNOIGKEKLGKLSSEDEKTEMEKAVEKIEWLESHODADIEDPKAK 619
QY 602 KKELEEVOPILSKLYGSAGPPPTGEEDPAE 632
DB 620 KKELEEVOPILSKLYGSAGPPPTGEEDPAE 650

```

RESULT 2
 US-08-441-139-16
 Sequence 16, Application US/08441139
 Patent No. 5773245
 GENERAL INFORMATION:
 APPLICANT: WILTRUP, DR. KARL D.
 TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
 TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
 STREET: 400 Garden City Plaza
 CITY: Garden City
 STATE: NY
 COUNTRY: USA
 ZIP: 11530
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/441.139
 FILING DATE: 15-MAY-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 08/089, 997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: DIGIGLIO, Frank S.
REGISTRATION NUMBER: 31, 346
REFERENCE/DOCKET NUMBER: 8646
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEO ID NO: 16
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-16

```

```

Query Match 98.2%; Score 3168; DB 1; Length 666;
Best local similarity 97.8%; Pred. No. 9,46-256;
Matches 617; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 2 EEDKKEDVGTAVGIDLTGTYSCVGFKNKGRVEIITANDGNRITPSYAFTPEGERLIGDA 61
DB 32 DEEKEDVGTAVGIDLTGTYSCVGFKNKGRVEIITANDGNRITPSYAFTPEGERLIGDA 91
QY 62 AKNQLTSPNENTVFDAKRLIGRTWMDPSVQODIKFLPKRYVEKTKPYIOVDIGGCGT 121
DB 92 AKNQLTSPNENTVFDAKRLIGRTWMDPSVQODIKFLPKRYVEKTKPYIOVDIGGCGT 151
QY 122 FAPBEISAMVLTAKMTAEAYLGKVTYHAYVTPAYFNDAROKRTKAGTACIAGLNMWRII 181
DB 152 FAPBEISAMVLTAKMTAEAYLGKVTYHAYVTPAYFNDAROKRTKAGTACIAGLNMWRII 211
QY 182 NEPTAAIAVGLDKREGKENILVFDLGGCTFDVSLITDNGVFEVATNGDTHLGEDPD 241
DB 212 NEPTAAIAVGLDKREGKENILVFDLGGCTFDVSLITDNGVFEVATNGDTHLGEDPD 271
QY 242 ORVMEHFIRLYKKTKGDKVKNRAVOKLRREVEKAKRALSQHOARIEIESFEYEGEDFS 301
DB 272 ORVMEHFIRLYKKTKGDKVKNRAVOKLRREVEKAKRALSQHOARIEIESFEYEGEDFS 331
QY 302 ETLTRAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLVGSTRIRKIOOLYKEFP 361
DB 332 ETLTRAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLVGSTRIRKIOOLYKEFP 391
QY 362 NGKEPSKGINPDEAVAYGAAGVLSGDODTGDVLVDVCPPLTGLGIEYGVWTKLIIPR 421
DB 392 NGKEPSKGINPDEAVAYGAAGVLSGDODTGDVLVDVCPPLTGLGIEYGVWTKLIIPR 451
QY 422 NTVPPTKKSQIFSTASDNPVTIIVYEGSERPLTKDNHLLCTFDLTIIPAPRGVPOIEV 481
DB 452 NTVPPTKKSQIFSTASDNPVTIIVYEGSERPLTKDNHLLCTFDLTIIPAPRGVPOIEV 511
QY 482 TFEIDVNGILRYTAEDEKGTGNKNKITITNDONRLTPEIEEMVNDAKFAEEDKKLKERI 541
DB 512 TFEIDVNGILRYTAEDEKGTGNKNKITITNDONRLTPEIEEMVNDAKFAEEDKKLKERI 571
QY 542 DTRNELESYAVSLKNOIGKEKLGKLSSEDEKTEMEKAVEKIEWLESHODADIEDPKAK 601
DB 572 DTRNELESYAVSLKNOIGKEKLGKLSSEDEKTEMEKAVEKIEWLESHODADIEDPKAK 631
QY 602 KKELEEVOPILSKLYGSAGPPPTGEEDPAE 632
DB 632 KKELEEVOPILSKLYGSAGPPPTGEEDPAE 662

```

RESULT 3
 US-08-441-139-2
 Sequence 2, Application US/08441139
 Patent No. 5773245